



QY 121 HOETPVYLGATAGMRLNMESEELADRVLDYVERSLSNYPDEFGARIIITGOEGAYGMI 180  
 DB 181 TINYLGKFSQKTRMFSIYPIYETNNQETFGALDLGASQVTFVPONOTIESPDNALQFR 240  
 QY 181 TINYLGKFSQKTRMFSIYPIYETNNQETFGALDLGASQVTFVPONOTIESPDNALQFR 240  
 DB 241 LYKGDYNYTHSHFLCYGKQDALMOKLAKDIOVASNEILRDPCHFGYKKVNVSDLYKTP 300  
 QY 241 LYKGDYNYTHSHFLCYGKQDALMOKLAKDIOVASNEILRDPCHFGYKKVNVSDLYKTP 300  
 DB 301 CTKRREMTLPFOQFEIQTGIGNYQOCHQSTILEFNISYCPYSOCAFNGIFLPLQGDGFGAF 360  
 QY 301 CTKRREMTLPFOQFEIQTGIGNYQOCHQSTILEFNISYCPYSOCAFNGIFLPLQGDGFGAF 360  
 DB 361 SAFYVPMKFLNLTSEKVSQEKYTEMMKRCQAPWEIKTSYAGYKEXYLSYCFSGTYIL 420  
 QY 361 SAFYVPMKFLNLTSEKVSQEKYTEMMKRCQAPWEIKTSYAGYKEXYLSYCFSGTYIL 420  
 DB 421 SLLQGYHFTADSWEHIFRIGKIQGSDAGWTGLGYMLNTNMIPEQPLSTPLSHSTYVFL 480  
 QY 421 SLLQGYHFTADSWEHIFRIGKIQGSDAGWTGLGYMLNTNMIPEQPLSTPLSHSTYVFL 480  
 DB 481 MYFLSLVLTVAIIIGLLIFHKPSYFKDMV 510  
 QY 481 MYFLSLVLTVAIIIGLLIFHKPSYFKDMV 510

## RESULT 2

ID W04264 standard; Protein: 510 AA.  
 AC W04264;  
 DT 29-MAY-1997 (first entry)  
 DE Human CD39 protein.  
 KW Human: lymphocyte activation marker; gene therapy;  
 KW ATP diphosphohydrolase; transgenic animal; pig; porcine; graft;  
 KW donor; transplantation; endothelial cell; prosthetic device;  
 KW platelet aggregation; inhibition; intravascular.  
 OS Homo sapiens.  
 PN W09630532-A1.  
 PD 03-OCT-1996.  
 PE 22-MAR-1996; E01270.  
 PR 24-MAR-1995; US-410371.  
 PR 12-FEB-1996; US-600383.  
 PA (NEMO-) NEW ENGLAND DEACONESS HOSPITAL.  
 PA (SANO-) SANDOZ LTD.  
 PI Bach FH, Robson S;  
 DR WPI; 96-455377/45.  
 DR N-PSDB: T33966.  
 PT Gene therapy of inflammatory or immunological stimulation of  
 PT platelet aggregation - using CD39 protein with ATP  
 PT d1:phospho:hydrolase activity; useful for preventing or alleviating  
 PT thrombotic condition in mammalian subject  
 PS Claim 3; Page 39; 65pp; English.  
 CC Non-human transgenic or somatic recombinant mammals, whose cells  
 CC contain a heterologous DNA encoding a polypeptide (especially human  
 CC CD39 protein) having ATP-diphosphohydrolase activity under cellular  
 CC activating conditions is claimed. In particular the animal is a pig  
 CC and its cells (or tissues or organs) can be used for transplantation.  
 CC DNA coding for human CD39 is also useful for genetically modifying a  
 CC mammalian cell to render it less susceptible to an inflammatory or  
 CC immunological stimulus and platelet aggregation. The modified cells  
 CC can be used to prevent or alleviate a thrombotic condition.  
 CC The present sequence is that of the human CD39 protein and was  
 CC disclosed in J Immunol. 153 (8) (1994) 3574-3584.  
 SO Sequence 510 AA;

Query Match 100.0%; Score 3806; DB 1; Length 510;  
 Best local similarity 100.0%; Pred. No. 0.00e+00;  
 Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MEDRESNWKTFCKSNIIAIIIGFSSIIAVIALNGLTQNKALPENWYKGYIADGSSHT 60  
 QY 1 MEDRESNWKTFCKSNIIAIIIGFSSIIAVIALNGLTQNKALPENWYKGYIADGSSHT 60

DB 61 SLTIYKPAEKENDTGVVQVECECRVKGPGISKFYQKVNIGIYITLDCMERAREVIPSQ 120  
 QY 61 SLTIYKPAEKENDTGVVQVECECRVKGPGISKFYQKVNIGIYITLDCMERAREVIPSQ 120  
 DB 121 HOETPVYLGATAGMRLNMESEELADRVLDYVERSLSNYPDEFGARIIITGOEGAYGMI 180  
 QY 121 HOETPVYLGATAGMRLNMESEELADRVLDYVERSLSNYPDEFGARIIITGOEGAYGMI 180  
 DB 181 TINYLGKFSQKTRMFSIYPIYETNNQETFGALDLGASQVTFVPONOTIESPDNALQFR 240  
 QY 181 TINYLGKFSQKTRMFSIYPIYETNNQETFGALDLGASQVTFVPONOTIESPDNALQFR 240  
 DB 241 LYKGDYNYTHSHFLCYGKQDALMOKLAKDIOVASNEILRDPCHFGYKKVNVSDLYKTP 300  
 QY 241 LYKGDYNYTHSHFLCYGKQDALMOKLAKDIOVASNEILRDPCHFGYKKVNVSDLYKTP 300  
 DB 301 CTKRREMTLPFOQFEIQTGIGNYQOCHQSTILEFNISYCPYSOCAFNGIFLPLQGDGFGAF 360  
 QY 301 CTKRREMTLPFOQFEIQTGIGNYQOCHQSTILEFNISYCPYSOCAFNGIFLPLQGDGFGAF 360  
 DB 361 SAFYVPMKFLNLTSEKVSQEKYTEMMKRCQAPWEIKTSYAGYKEXYLSYCFSGTYIL 420  
 QY 361 SAFYVPMKFLNLTSEKVSQEKYTEMMKRCQAPWEIKTSYAGYKEXYLSYCFSGTYIL 420  
 DB 421 SLLQGYHFTADSWEHIFRIGKIQGSDAGWTGLGYMLNTNMIPEQPLSTPLSHSTYVFL 480  
 QY 421 SLLQGYHFTADSWEHIFRIGKIQGSDAGWTGLGYMLNTNMIPEQPLSTPLSHSTYVFL 480  
 DB 481 MYFLSLVLTVAIIIGLLIFHKPSYFKDMV 510  
 QY 481 MYFLSLVLTVAIIIGLLIFHKPSYFKDMV 510

## RESULT 3

ID W67846 standard; Protein: 323 AA.  
 AC W67846;  
 DT 25-MAR-1999 (first entry)  
 DE Human secreted protein encoded by gene 40 clone HTOEU03.  
 KW Human: secreted protein; fusion protein; gene therapy; protein therapy;  
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;  
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
 KW inflammation; ischaemic shock; Alzheimer's disease; osteoclast; thymus;  
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
 OS Homo sapiens.  
 PN W09842738-A1.  
 PD 01-OCT-1998.  
 PE 19-MAR-1998; U05311.  
 PR 30-MAY-1997; US-050937.  
 PR 21-MAR-1997; US-041276.  
 PR 21-MAR-1997; US-041277.  
 PR 21-MAR-1997; US-041281.  
 PR 21-MAR-1997; US-042344.  
 PR 30-MAY-1997; US-048069.  
 PR 30-MAY-1997; US-048094.  
 PR 30-MAY-1997; US-048095.  
 PR 30-MAY-1997; US-048096.  
 PR 30-MAY-1997; US-048099.  
 PR 30-MAY-1997; US-048131.  
 PR 30-MAY-1997; US-048135.  
 PR 30-MAY-1997; US-048154.  
 PR 30-MAY-1997; US-048160.  
 PR 30-MAY-1997; US-048186.  
 PR 30-MAY-1997; US-048187.  
 PR 30-MAY-1997; US-048188.  
 PR 30-MAY-1997; US-048350.  
 PR 30-MAY-1997; US-048351.  
 PR 30-MAY-1997; US-048352.  
 PR 30-MAY-1997; US-048355.  
 PR 05-AUG-1997; US-054804.  
 PA (HUMA-) HUMAN GENOME SCI INC.



FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	243	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	269	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	285	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	312	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	319	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	344	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	358	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	378	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	386	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	465	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	473	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	473	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	494	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	555	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	593	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	602	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	609	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	621	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	724	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	736	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	739	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	786	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	841	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	924	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	934	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	1017	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	1054	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	1127	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	1147	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	/note-	"corresponds to stop codon in DNA sequence"
FT	1178	/note-	"corresponds to stop codon in DNA sequence"
PN	W09632476-A1.	/note-	"corresponds to stop codon in DNA sequence"
PD	17-OCT-1996.	/note-	"corresponds to stop codon in DNA sequence"
PE	12-APR-1996: CA0214.	/note-	"corresponds to stop codon in DNA sequence"
PR	13-APR-1995: US-421701.	/note-	"corresponds to stop codon in DNA sequence"
PA	(MOUNT ) MOUNT SINAI HOSPITAL CORP.	/note-	"corresponds to stop codon in DNA sequence"
PI	Labes M, Lozano A, Roach A, Roder J;	/note-	"corresponds to stop codon in DNA sequence"
DR	WPI; 96-477127/47.	/note-	"corresponds to stop codon in DNA sequence"
DR	N-PSDB: T38484.	/note-	"corresponds to stop codon in DNA sequence"
PT	Assay for substance that modulates response of neuronal cells - and	/note-	"corresponds to stop codon in DNA sequence"
PT	neurite growth associated protein, Petrin, useful in conditions	/note-	"corresponds to stop codon in DNA sequence"
PT	involving nerve damage resulting from traumatic injury, stroke or	/note-	"corresponds to stop codon in DNA sequence"
PT	CNS degenerative disorders	/note-	"corresponds to stop codon in DNA sequence"

PS	Claim 9; Page 57-61; 119pp; English.
CC	Rat petrin (W043326) is a protein involved in modulating neurite growth inhibition. The amino sequence was deduced from a cDNA clone (T834484) derived from an adult rat brain cDNA library; no coding sequence was indicated. Petrin is a new member of the CC protein phosphatase 2C family, and is expressed in neurons in brain tissue, partic. in the Purkinje cells of the cerebellum. Petrin, CC and antibodies raised against it, can be used to modulate neurite growth and axonal regeneration.
CC	Growth and axonal regeneration.
SQ	Sequence 1196 AA;
DB	Query Match 2.8%; Score 106; DB 1; Length 1196; Best Local Similarity 31.4%; Pred. No. 1.51e+01; Matches 16; Conservative 15; Mismatches 18; Indels 2; Gaps 2.
YQ	503 KYEQQSHCVQLFQGSFLPISCPILRFSIFCGDGVQSFFGFEVL-FL 552 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  321 NMQQGHSLELFNMTSYCPYSQCARN-GIPLPLQGDGCARSATFYFWKFL 370
RESULT	7
ID	W04339 standard; Protein; 20 AA.
AC	W04339;
DR	29-DEC-1996 (first entry)
DE	ATP diaphosphohydrolase fragment from pig pancreas.
KM	ATP diaphosphohydrolase; ATPases; bovine aorta; pig pancreas;
KW	Apyrase; CD39; lymphoid cell activation antigen; enzyme;
KW	platelet aggregation; thrombogenicity; anti-haemostatic.
OS	Sus scrofa.
PN	W069632471-A2.
PD	17-OCT-1996.
PF	10-APR-1996; CA0223.
PR	10-APR-1995; US-419204.
PA	(UYSH ) UNIV SHERBROOKE.
PI	Beaudoin AR, Sevigny J; WPI: 96-477122/47.
PT	Isolated ATP di:phospho:hydrolase enzymes - have anti-haemostatic activity, useful for reducing platelet aggregation and PR thrombogenicity
PT	thrombogenicity
PS	Claim 3; Page 46; 60pp; English.
CC	The bovine aorta (W04335 to W04338 and W04340) and porcine pancreatic (W04339) ATPases have been partially sequenced.
CC	The sequences have been found to be highly homologous to a human lymphoid cell activation antigen designated CD39 (Maliszewski et al. (1994). J. Immunol.: 3574-3583). The complete sequences of the CC ATPases types I and II have not been obtained yet. Assuming that CC the CD39 gene product is an ATPase type II, the use of CD39 in the reduction of platelet aggregation and of thrombogenicity CC may be contemplated, as well as a process of making ATPases CC using the CD39 sequence (t38516).
SQ	Sequence 20 AA;
DB	Query Match 2.7%; Score 104; DB 1; Length 20; Best Local Similarity 78.9%; Pred. No. 2.06e+01; Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0.
YQ	2 SDPOETFGALDIGASTOV 20 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  203 TNNQETFGALDIGASTOV 221
RESULT	8
ID	W20877 standard; Protein; 95 AA.
AC	W20877.
DR	18-JUL-1997 (first entry)
DE	H. pylori cytoplasmic protein, 13ee10215orf82.
KM	Cytoplasmic; vaccine; prevention; treatment; envelope;
KW	identification; binding compound; bacterium; life cycle; activator;
KW	bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
OS	Helicobacter pylori.
PN	W09640893-Al.
PD	19-DEC-1996.



PF 06-JUN-1996; U09122.  
PR 07-JUN-1995; US-487032.  
PR 01-APR-1996; US-630405.  
PA (ASTR ) ASTRA AB.  
PI Berglindh O, Smith D, Møllgaard BL;  
DR WPI; 97-052306/05.  
DR N-PSDB; T68130.  
PT Helicobacter pylori nucleic acid sequences and related polypeptide(s) - useful for vaccines to treat or prevent H. pylori infection, and to detect Helicobacter.  
PS Claim 61; Page 1278-79; 1481pp; English.  
CC This sequence represents a H. pylori cytoplasmic protein.  
CC The protein may be used in a vaccine to prevent or treat H. pylori infection or to identify H. pylori polypeptide binding compounds, CC useful as potential H. pylori life cycle activators or inhibitors.  
CC The genomic sequence of H. pylori (ATCC 55679) was determined from overlapping contigs generated by mechanically shearing the bacterial DNA. The sequences were analysed for ORF of at least 180 nucleotides, CC and the predicted coding regions defined by computer evaluation. To CC identify likely H. pylori antigens for vaccine development, the amino CC acid sequences predicted from various ORF were analysed for significant CC homology to other known or exported membrane proteins. Having identified CC and determined the sequences of interest, particular regions can be CC isolated from H. pylori by PCR amplification for recombinant polypeptide CC production, e.g. in E. coli hosts.  
SQ Sequence 95 AA;

Query Match 2.7%; Score 103; DB 1; Length 95;  
Best Local Similarity 34.0%; Pred. No. 2.41e+01;  
Matches 18; Conservative 16; Mismatches 14; Indels 5; Gaps 5;

Db 24 GYM-VSLGALKRPPCRREFLKALFAIF-YHAVNNPLTQCPRPHQREFSS 74  
QY 286 GYKRVNVSDLYKPPCTKRFEM-TLPQOEFIGIGNY-QQCH-QSILELFT 335

RESULT 9  
ID R85880 standard; Protein; 439 AA.  
AC R85880;  
DT 13-SEP-1996 (first entry)  
DE WD-40 domain-contg. YC07 protein.  
KW WP40 repeat region; beta-transducin; protein-protein interaction; drug; intracellular signalling; protein kinase C; homology; motif; modulator; KM receptors of activated protein kinase; enzyme activity; isozyme; human.  
OS Synthetic.  
PN W09521252-A2.  
PD 10-AUG-1995.  
PR 31-JAN-1995; U01210.  
PR 01-FEB-1994; US-190802.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Mochly-Rosen D, Ron D;  
DR WPI; 95-283772/37.  
PT New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the PT activity of a protein, eg. protein kinase C, which interacts with a PT protein contg. a WD-40 region.  
PS Example 5; Page 166-168; 351pp; English.  
CC Proteins R85851-92 are protein which contain at least one WD-40 (also CC called beta-transducin homologous) amino acid repeat motifs. The WD-40 CC regions are involved in protein-protein interactions between proteins CC involved in intracellular signalling. An example of such an interaction CC is between protein kinase C and receptors of activated protein kinase CC (RACK), esp. RACK-1 (R85850). Proteins R85851-82 were isolated based on CC homology with beta-transducin, whereas proteins R85882-92 were isolated CC based on homology with the WD-40 consensus sequence (R85893). The CC proteins were used to construct the peptides R84928-R85063 and CC R85786-R85842. The peptides can be used to identify target proteins CC contg. WD-40 motifs, as modulators of enzyme esp. isozyme, activity of CC proteins involved in protein-protein interaction and to screen for drugs CC that will affect protein-protein interaction involving WD-40 domains.  
SQ Sequence 439 AA;

Query Match 2.7%; Score 102; DB 1; Length 439;  
Best Local Similarity 26.7%; Pred. No. 2.81e+01;

Matches 16; Conservative 16; Mismatches 25; Indels 3; Gaps 3;

Db 180 AYPDGRVATASEDKIKWDITSGCLATFEHTSSVAVQAKRQWFS- SLDGT 238  
QY 160 PFDEGARITIGOEAGV-GW-ITINYLAKFSOKTRWFSIVPEYTNQETFGALDLOGA 217

RESULT 10  
ID W55029 standard; Protein; 481 AA.  
AC W55029;  
DT 17-SEP-1998 (first entry)  
DE G-protein coupled receptor, short form.  
KW G-protein coupled receptor; gene therapy; abnormality detection; KM short form; human.  
OS Homo sapiens.  
PN EP-845529-A2.  
PD 03-JUN-1998.  
PR 27-OCT-1997; 308562.  
PR 29-OCT-1996; JP-286823.  
PA (TAKE ) TAKEDA CHEM IND LTD.  
PI Fukusumi S, Hinuma S, Kawamata Y;  
DR WPI; 98-288746/26.  
DR N-PSDB; V27174.  
PT New human G-protein coupled receptor protein - and corresponding PT DNA, ligands, antibodies, etc  
PS Claim 1; Page 42-44; 65pp; English.  
CC This sequence represents a human G-protein coupled receptor of the CC invention. The protein or cells expressing the DNA encoding it can be CC used to screen for agonists or antagonists of the receptor, which can be CC used as drugs for treating various diseases (none disclosed). The DNA can CC also be used for practice drug design based on comparisons with CC structurally analogous ligands and receptors. DNA encoding the protein CC can be used for gene therapy for diseases caused by a deficiency of the CC receptor. The DNA can also be used to detect abnormalities in the gene CC encoding the receptor. The protein or fragment can be used to determine CC levels of receptor ligands in vivo. The antibody can be used in assays to CC detect the protein.  
SQ Sequence 481 AA;

Query Match 2.7%; Score 101; DB 1; Length 481;  
Best Local Similarity 35.1%; Pred. No. 3.28e+01;  
Matches 13; Conservative 14; Mismatches 8; Indels 2; Gaps 2;

Db 114 LQIONPLX-PYTESYSATMALALVFPVAVGIVGNL 149  
QY 462 IPAEQPLSTPLSHSTY-VFLMVLFSLVLTVAIIGLL 497

RESULT 11  
ID W55030 standard; Protein; 542 AA.  
AC W55030;  
DT 17-SEP-1998 (first entry)  
DE G-protein coupled receptor, long form.  
KW G-protein coupled receptor; gene therapy; abnormality detection; KM long form; human.  
OS Homo sapiens.  
PN EP-845529-A2.  
PD 03-JUN-1998.  
PR 27-OCT-1997; 308562.  
PR 29-OCT-1996; JP-286823.  
PA (TAKE ) TAKEDA CHEM IND LTD.  
PI Fukusumi S, Hinuma S, Kawamata Y;  
DR WPI; 98-288746/26.  
DR N-PSDB; V27175.  
PT New human G-protein coupled receptor protein - and corresponding PT DNA, ligands, antibodies, etc  
PS Claim 2; Page 45-47; 65pp; English.  
CC This sequence represents a human G-protein coupled receptor of the CC invention. The protein, or cells expressing the DNA encoding it can be CC used to screen for agonists or antagonists of the receptor, which can be CC used as drugs for treating various diseases (none disclosed). The DNA can CC also be used for practice drug design based on comparisons with CC structurally analogous ligands and receptors. DNA encoding the protein















CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Frommer, William S.  
CC REGISTRATION NUMBER: 25,506  
CC REFERENCE/DOCKET NUMBER: 454310-2540  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 840-3333  
CC TELEFAX: (212) 840-0712  
CC TELEX: 425066 CURTMS  
CC INFORMATION FOR SEQ ID NO: 12:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 459 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC FRAGMENT TYPE: internal  
CC SEQUENCE 459 AA; 52308 MW; 1156694 CN;  
SO  
Query Match 2.5%; Score 96; DB 1; Length 459;  
Best Local Similarity 32.6%; Pred. No. 3.50e+01;  
Matches 14; Conservative 12; Mismatches 16; Indels 1; Gaps 1;  
Db 344 CYSRENKILQVLKQIPINGDIIISGCLNHPGLVNIQNKIDI 386  
QY 255 CYGDOALWQKLAKIDQVANSNELLRDPCE-HPGYKKVNVSDL 296  
RESULT 7  
ID US-08-923-772-2 STANDARD; PRT; 523 AA.  
XX xxxxxx  
AC  
DT  
DE Sequence 2, Application US/08923772  
XX  
CC Sequence 2, Application US/08923772  
CC Patent No. 5972651  
CC GENERAL INFORMATION:  
CC APPLICANT: Black, Michael T.  
CC TITLE OF INVENTION: NOVEL FIB  
CC NUMBER OF SEQUENCES: 2  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Dechert Price & Rhoads  
CC STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
CC CITY: Philadelphia  
CC STATE: PA  
CC COUNTRY: US  
CC ZIP: 19103  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Diskette  
CC OPERATING SYSTEM: DOS  
CC SOFTWARE: FastSeq for Windows Version 2.0  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/923,772  
CC FILING DATE:  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC ATTORNEY/AGENT INFORMATION:  
CC FILING DATE:  
CC NAME: Dickinson, Todd O  
CC REGISTRATION NUMBER: 28,354  
CC REFERENCE/DOCKET NUMBER: GMI0080  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 215-994-2252  
CC TELEFAX: 215-994-2222  
CC TELEX:  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 523 amino acids  
CC

CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC SEQUENCE 523 AA; 57775 MW; 1280224 CN;  
SO  
Query Match 2.5%; Score 94; DB 2; Length 523;  
Best Local Similarity 20.7%; Pred. No. 4.72e+01;  
Matches 29; Conservative 40; Mismatches 64; Indels 7; Gaps 7;  
Db 196 LQYDELLMNEIDVQVLAQPNELLVYDAMIGQEAANYAREFNAQLEV-TGVIIITKIDGD 254  
QY 27 IAVIALAVGLTQNKALPENVKYGIYLDGSSHTSLYIK-WPAKENDTGVMHVEECR 85  
Db 255 TRG-GAALSVRHITGKPIKFTGTGKRTDIEFTHDPDMSRI-LGMDMLTLIEKASQGY 312  
QY 86 VKGPGISKRFVQKYNELGILYIDOMERAREVIV-RSQHGTPEYLGATAGMRLIRSESEL 144  
Db 313 DECKALEMAEK-WRENTDF 331  
QY 145 AD-RVIDVERSLSNYPDF 163  
RESULT 8  
ID US-08-466-589-6 STANDARD; PRT; 627 AA.  
XX xxxxxx  
AC  
DT  
DE Sequence 6, Application US/08466589  
XX  
CC Sequence 6, Application US/08466589  
CC Patent No. 5837489  
CC GENERAL INFORMATION:  
CC APPLICANT: Elliot, Kathryn J.  
CC APPLICANT: Ellis, Steven B.  
CC APPLICANT: Harpold, Michael M.  
CC TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE  
CC TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME  
CC NUMBER OF SEQUENCES: 12  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Brown, Martin, Haller & McClaim  
CC STREET: 1660 Union Street  
CC CITY: San Diego  
CC STATE: CA  
CC COUNTRY: US  
CC ZIP: 92101-2926  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Diskette  
CC OPERATING SYSTEM: DOS  
CC SOFTWARE: FastSeq Version 1.5  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/466,589  
CC FILING DATE: June 5, 1995  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/028,031  
CC FILING DATE: March 8, 1993  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Seldman, Stephanie L  
CC REGISTRATION NUMBER: 33,779  
CC REFERENCE/DOCKET NUMBER: 6362-9950  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 619-238-0999  
CC TELEFAX: 619-238-0062  
CC TELEX:  
CC INFORMATION FOR SEQ ID NO: 6:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 627 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: unknown  
CC

CC MOLECULE TYPE: protein  
SQ SEQUENCE 627 AA; 69767 MW; 2152270 CN;  
Query Match 2.5%; Score 97; DB 2; Length 627;  
Best Local Similarity 26.9%; Pred. No. 3.01e+01;  
Matches 14; Conservative 22; Mismatches 12; Indels 4; Gaps 4;  
Db 287 GELLITETIIPPT-SLVIPLI-GEYLLFTMTFVTLST-AITVFVLNVHHRSP 335  
Y 453 GYMLNTLNMIPAEQPLSTPLSHSTYV-FLMWFLSLVLTVAIIGLLIFHKPS 503  
RESULT 9  
ID US-08-700-636-6 STANDARD; PRT; 627 AA.  
AC xxxxxx  
XX  
DT  
DE Sequence 6, Application US/08700636  
CC  
CC Sequence 6, Application US/08700636  
CC Patent No. 5910582  
CC GENERAL INFORMATION:  
CC APPLICANT: Elliot, Kathryn J.  
CC APPLICANT: Ellis, Steven B.  
CC APPLICANT: Harpold, Michael M.  
CC TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE  
CC TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME  
CC NUMBER OF SEQUENCES: 12  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
CC STREET: 444 South Flower Street, Suite 2000  
CC CITY: Los Angeles  
CC STATE: CA  
CC COUNTRY: USA  
CC ZIP: 90071  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/700,636  
CC FILING DATE: 16-JUL-1996  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/028,031  
CC FILING DATE: 08-MAR-1993  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Reiter, Stephen E.  
CC REGISTRATION NUMBER: 31,192  
CC REFERENCE/DOCKET NUMBER: P41 9368  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 619-546-4737  
CC TELEFAX: 619-546-9392  
CC INFORMATION FOR SEQ ID NO: 6:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 627 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: unknown  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 627 AA; 69767 MW; 2152270 CN;  
Query Match 2.5%; Score 97; DB 2; Length 627;  
Best Local Similarity 26.9%; Pred. No. 3.01e+01;  
Matches 14; Conservative 22; Mismatches 12; Indels 4; Gaps 4;  
Db 287 GELLITETIIPPT-SLVIPLI-GEYLLFTMTFVTLST-AITVFVLNVHHRSP 335  
Y 453 GYMLNTLNMIPAEQPLSTPLSHSTYV-FLMWFLSLVLTVAIIGLLIFHKPS 503

RESULT 10  
ID US-08-480-662-2 STANDARD; PRT; 971 AA.  
AC xxxxxx  
XX  
DT  
DE Sequence 2, Application US/08480662  
XX  
CC Sequence 2, Application US/08480662  
CC Patent No. 5759782  
CC GENERAL INFORMATION:  
CC APPLICANT: Pastan, Ira  
CC APPLICANT: Brinkmann, Ulrich  
CC TITLE OF INVENTION: CELLULAR APOPTOSIS SUSCEPTIBILITY PROTEIN (CSP) AND AN  
CC NUMBER OF SEQUENCES: 10  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Knobbe, Martens, Olson and Bear  
CC STREET: 620 Newport Center Drive 16th Floor  
CC CITY: Newport Beach  
CC STATE: CA  
CC COUNTRY: USA  
CC ZIP: 92660  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Diskette  
CC COMPUTER: IBM compatible  
CC OPERATING SYSTEM: DOS  
CC SOFTWARE: FASTSEQ Version 1.5  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/480,662  
CC FILING DATE: 07-JUN-1995  
CC CLASSIFICATION: 530  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER:  
CC FILING DATE:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Israelsen, Ned A  
CC REGISTRATION NUMBER: 29,655  
CC REFERENCE/DOCKET NUMBER: NIH112,001A  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 619-235-8550  
CC TELEFAX: 619-235-0176  
CC TELEX:  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 971 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC HYPOTHETICAL: NO  
CC ANTI-SENSE: NO  
CC FRAGMENT TYPE: N-terminal  
CC ORIGINAL SOURCE:  
SQ SEQUENCE 971 AA; 110313 MW; 4888742 CN;  
Query Match 2.5%; Score 94; DB 1; Length 971;  
Best Local Similarity 28.8%; Pred. No. 4.72e+01;  
Matches 17; Conservative 17; Mismatches 21; Indels 4; Gaps 4;  
Db 179 AFALPLTNLEKATIELCSTHANDASA-LRIFFSLILSKLFYSL-NFODLPETWEGNM 235  
Y 453 GYMLNTLNMIPAEQPLSTPLSHSTYVFLMWFLS-LVLTVAIIGLLIFHKPSYFWK-DM 509  
RESULT 11  
ID PCT-US96-09927-2 STANDARD; PRT; 971 AA.  
AC xxxxxx  
XX  
DT  
XX

```
DE      Sequence 2, Application PC/TUS9609927
CC
XX
CC      Sequence 2, Application PC/TUS9609927
CC      GENERAL INFORMATION:
CC      APPLICANT: The United States, As Represented by the
CC      APPLICANT: Secretary, Department of Health and Human
CC      APPLICANT: Services
CC      TITLE OF INVENTION: CELLULAR APOPTOSIS
CC      NUMBER OF SEQUENCES: 14
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Knobbie, Martens, Olson and Bear
CC      STREET: 620 Newport Center Drive 16th Floor
CC      CITY: Newport Beach
CC      STATE: CA
CC      COUNTRY: USA
CC      ZIP: 92660
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Diskette
CC      COMPUTER: IBM Compatible
CC      OPERATING SYSTEM: DOS
CC      SOFTWARE: FASTSEQ Version 1.5
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: PCT/US96/09927
CC      FILING DATE: 07-JUN-1995
CC      CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US08/480662
CC      FILING DATE: 07-JUN-1995
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Israelsen, Ned A
CC      REGISTRATION NUMBER: 29,655
CC      REFERENCE/DOCKET NUMBER: NIH112.0010PC
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: 619-235-8550
CC      TELEFAX: 619-235-0176
CC      TELEX:
CC      INFORMATION FOR SEQ ID NO: 2:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 971 amino acids
CC      TYPE: amino acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: peptide
CC      HYPOTHETICAL: NO
CC      ANTI-SENSE: NO
CC      FRAGMENT TYPE: N-terminal
CC      ORIGINAL SOURCE:
CC      SEQUENCE 971 AA; 110313 MW; 4888742 CN;
S0
DB      Query Match 2.5%; Score 94; DB 3; Length 971;
DB      Best Local Similarity 28.8%; Pred. No. 4.72e+01;
QY      Matches 17; Conservative 17; Mismatches 21; Indels 4; Gaps 4
DB      179 AFAPLPLTNLFKATITELCTSTHANDASA-LRIIFSSILILSKFLYSI-NFODLPETWEGNM 235
QY      453 GYMILNLTWNIPAEOPSLPISHSHTVFLMWIFS-LVLEPTVAIGILLIFHKPSYFWK-DM 509
::| ||||: | : :: | |||| | : : | : | : |
RESULT 12
ID      US-07-582-945-2 STANDARD; PRT; 1285 AA.
XX
AC      xxxxxx
XX
DT
XX
DE      Sequence 2, Application US/07582945
CC
CC      Sequence 2, Application US/07582945
CC      Patent No. 5369019
CC      GENERAL INFORMATION:
CC      APPLICANT: TAKKER FOGED, Niels
CC      APPLICANT: PETERSON, Svend
```

```

CC      TITLE OF INVENTION:  A PASTEURELLA VACCINE
CC      NUMBER OF SEQUENCES:  2
CC      CORRESPONDENCE ADDRESSES:
CC      ADDRESSEE:  Foley & Lardner
CC      STREET:  1800 Diagonal Road, Suite 500
CC      CITY:  Alexandria
CC      STATE:  VA
CC      COUNTRY:  USA
CC      ZIP:  22313-0299
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE:  Floppy disk
CC      COMPUTER:  IBM PC compatible
CC      OPERATING SYSTEM:  PC-DOS/MS-DOOS
CC      SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER:  US/07/582,945
CC      FILING DATE:  19901012
CC      CLASSIFICATION:  424
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME:  BENT, Stephen A.
CC      REGISTRATION NUMBER:  29,768
CC      REFERENCE/DOCKET NUMBER:  30307/112 PLVI
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE:  (703)836-9300
CC      TELEFAX:  (703)683-4109
CC      TELEX:  899149
CC      INFORMATION FOR SEQ ID NO:  2:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH:  1285 amino acids
CC      TYPE:  AMINO ACID
CC      TOPOLOGY:  linear
CC      MOLECULE TYPE:  protein
CC      SEQUENCE  1285 AA; 146565 MW; 8485412 CN;
SQ
Query Match          2.5%;  Score 95;  DB 1;  Length 1285;
Best Local Similarity 30.6%;  Pred. NO. 4.07e+01;
Matches  13;  Conservative  13;  Mismatches  26;  Indels  4;  Gaps  4;
Db      823 PYELPTFFTFMSLQVQSDLG-FEQAFATRRFFNTLVSDRLSLMENTMLLTSEFDYTPMD 881
QY      339 PYSQCAFNGI-FLPPLQGDGFASFAFYVMKFLN-LTSEKVS-QEYTEMKKKCAQPWE 395
        ||  :  :  :  |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      882 AI 883
QY      396 EI 397
RESULT  13
ID      US-08-453-141-2          STANDARD:          PRT;      1285 AA.
XX      AC      xxxxxx
XX      DT
XX
DE      Sequence 2, Application US/08453141
XX
CC      Sequence 2, Application US/08453141
CC      Patent No. 5885589
CC      GENERAL INFORMATION:
CC      APPLICANT:  FOGED, Niels T.
CC      APPLICANT:  PETERSEN, Svend
CC      TITLE OF INVENTION:  PASTEURELLA VACCINE
CC      NUMBER OF SEQUENCES:  2
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE:  Foley & Lardner
CC      STREET:  3000 K Street, N.W., Suite 500
CC      CITY:  Washington, D.C.
CC      COUNTRY:  USA
CC      ZIP:  20007-5109
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE:  Floppy disk
CC      COMPUTER:  IBM PC compatible
CC      OPERATING SYSTEM:  PC-DOS/MS-DOOS

```



	Query Match	2.4%	Score 93;	DB 2;	Length 175;
	Best Local Similarity	29.1%	Pred. No.	5,48e+01;	
	Matches	16;	Conservative	13;	Mismatches 24; Indels 2; Gaps 2;
D6	65	LNLGQALFOEIVITLD-FSTPGAAVAANVFLSDKWESEHNHSGGVLP	118		
OY	370	LNLTSERKSQEAVTDMKKFCNQPEELKTISAG-VKENIYSECFSTIIILSL	423		

Search completed: Fri May 5 09:00:32 2000  
Job time : 81 secs.



CC MOLECULE TYPE: Protein  
SQ SEQUENCE 510 AA; 57964 MW; 1469467 CN;  
Query Match 100.0%; Score 3806; DB 12; Length 510;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MEDTKESNVKTFCSKNILALIGFSSIIAVIALAVGLTQNKALPENVKYGIYLDAGSSHT 60  
1 MEDTKESNVKTFCSKNILALIGFSSIIAVIALAVGLTQNKALPENVKYGIYLDAGSSHT 60  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 61 SLIYIKWPAEKENDGVYHQAQVEECVKGPGISKFOVKNEIGIYITDCMERAREVIIPRSQ 120  
61 SLIYIKWPAEKENDGVYHQAQVEECVKGPGISKFOVKNEIGIYITDCMERAREVIIPRSQ 120  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 121 HOETPVYLGATAGMRLNRESEELADRVLDYVERSLSNYPDFOGARITITGDEGAYGWI 180  
121 HOETPVYLGATAGMRLNRESEELADRVLDYVERSLSNYPDFOGARITITGDEGAYGWI 180  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 181 TINYLIGKFSQKTRMFSTIVPYETNNQETFGALDAGASTQVTFVPONOTIESPDNALQFR 240  
181 TINYLIGKFSQKTRMFSTIVPYETNNQETFGALDAGASTQVTFVPONOTIESPDNALQFR 240  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 241 LYGKDYNYTTHSFILCYGKDQALMOKLADIQVANSNEILRDPCHFPGYKKVNVSDLYKTP 300  
241 LYGKDYNYTTHSFILCYGKDQALMOKLADIQVANSNEILRDPCHFPGYKKVNVSDLYKTP 300  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 301 CTKRREMTLPFOQFEIIGIGNYQOCHOSTLEFNTSYCYSCAFNGIFLPLQGDGAF 360  
301 CTKRREMTLPFOQFEIIGIGNYQOCHOSTLEFNTSYCYSCAFNGIFLPLQGDGAF 360  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 361 SAFYFVMKFLNLTSEKVSQEKYTEMKKRFCAOPMEIKTSYAGVKEKYLSEYCFSGTYIL 420  
361 SAFYFVMKFLNLTSEKVSQEKYTEMKKRFCAOPMEIKTSYAGVKEKYLSEYCFSGTYIL 420  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 421 SLLQGYHFTADSWEHIFIKIGQSDAGWTLGYMLNTNMIIPAOPLSTPLSHSTYVFL 480  
421 SLLQGYHFTADSWEHIFIKIGQSDAGWTLGYMLNTNMIIPAOPLSTPLSHSTYVFL 480  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 481 MYLSVLFTVAITGLIFHKPSYFWKDMV 510  
481 MYLSVLFTVAITGLIFHKPSYFWKDMV 510  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 2  
ID US-09-374-586-1 STANDARD; PRT: 510 AA.  
AC xxxxxx  
DT  
XX  
XX Sequence 1, Application US/09374586  
DE  
XX Sequence 1, Application US/09374586  
CC Sequence 1, Application US/09374586  
CC GENERAL INFORMATION:  
CC APPLICANT: Pinsky, David J.  
CC TITLE OF INVENTION: CD39/ECTO-ADPASE AS A TREATMENT FOR THROMBOTIC AND  
CC TITLE OF INVENTION: ISCHEMIC DISORDERS  
CC FILE REFERENCE: 0575/59167  
CC CURRENT APPLICATION NUMBER: US/09/374,586  
CC CURRENT FILING DATE: 1999-08-13  
CC NUMBER OF SEQ ID NOS: 2  
CC SOFTWARE: PatentIn Ver. 2.1  
CC SEQ ID NO 1  
CC LENGTH: 510  
CC TYPE: PRT  
CC ORGANISM: HOMO-SAPIEN  
CC SEQUENCE 510 AA; 57964 MW; 1469467 CN;  
SQ  
Query Match 100.0%; Score 3806; DB 25; Length 510;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MEDTKESNVKTFCSKNILALIGFSSIIAVIALAVGLTQNKALPENVKYGIYLDAGSSHT 60  
1 MEDTKESNVKTFCSKNILALIGFSSIIAVIALAVGLTQNKALPENVKYGIYLDAGSSHT 60  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 61 SLIYIKWPAEKENDGVYHQAQVEECVKGPGISKFOVKNEIGIYITDCMERAREVIIPRSQ 120  
61 SLIYIKWPAEKENDGVYHQAQVEECVKGPGISKFOVKNEIGIYITDCMERAREVIIPRSQ 120  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 121 HOETPVYLGATAGMRLNRESEELADRVLDYVERSLSNYPDFOGARITITGDEGAYGWI 180  
121 HOETPVYLGATAGMRLNRESEELADRVLDYVERSLSNYPDFOGARITITGDEGAYGWI 180  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 181 TINYLIGKFSQKTRMFSTIVPYETNNQETFGALDAGASTQVTFVPONOTIESPDNALQFR 240  
181 TINYLIGKFSQKTRMFSTIVPYETNNQETFGALDAGASTQVTFVPONOTIESPDNALQFR 240  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 241 LYGKDYNYTTHSFILCYGKDQALMOKLADIQVANSNEILRDPCHFPGYKKVNVSDLYKTP 300  
241 LYGKDYNYTTHSFILCYGKDQALMOKLADIQVANSNEILRDPCHFPGYKKVNVSDLYKTP 300  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 301 CTKRREMTLPFOQFEIIGIGNYQOCHOSTLEFNTSYCYSCAFNGIFLPLQGDGAF 360  
301 CTKRREMTLPFOQFEIIGIGNYQOCHOSTLEFNTSYCYSCAFNGIFLPLQGDGAF 360  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 361 SAFYFVMKFLNLTSEKVSQEKYTEMKKRFCAOPMEIKTSYAGVKEKYLSEYCFSGTYIL 420  
361 SAFYFVMKFLNLTSEKVSQEKYTEMKKRFCAOPMEIKTSYAGVKEKYLSEYCFSGTYIL 420  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 421 SLLQGYHFTADSWEHIFIKIGQSDAGWTLGYMLNTNMIIPAOPLSTPLSHSTYVFL 480  
421 SLLQGYHFTADSWEHIFIKIGQSDAGWTLGYMLNTNMIIPAOPLSTPLSHSTYVFL 480  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 481 MYLSVLFTVAITGLIFHKPSYFWKDMV 510  
481 MYLSVLFTVAITGLIFHKPSYFWKDMV 510  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 3  
ID US-08-930-921-1 STANDARD; PRT: 510 AA.  
AC xxxxxx  
DT  
XX  
XX Sequence 1, Application US/08930921B  
DE  
XX Sequence 1, Application US/08930921B  
CC GENERAL INFORMATION:  
CC APPLICANT: BEAUDOIN, Adrien R.  
CC APPLICANT: SEVIGNY, Jean  
CC TITLE OF INVENTION: ATP-DIPHOSPHOHYDROLASES, PROCESS OF PURIFICATION  
CC TITLE OF INVENTION: THEROOF AND PROCES OF PRODUCING THEROOF BY RECOMBINANT  
CC TITLE OF INVENTION: TECHNOLOGY  
CC FILE REFERENCE: ATP-DIPHOSPHOHYDROLASES-BEAUDOIN  
CC CURRENT APPLICATION NUMBER: US/08/930,921B  
CC EARLIER FILING DATE: 1998-01-02  
CC EARLIER APPLICATION NUMBER: PCT/CA96/00223  
CC NUMBER OF SEQ ID NOS: 8  
CC SOFTWARE: PatentIn Ver. 2.0  
CC SEQ ID NO 1  
CC LENGTH: 510  
CC TYPE: PRT  
CC ORGANISM: Unknown  
CC OTHER INFORMATION: Description of Unknown Organism: unknown  
CC SEQUENCE 510 AA; 57964 MW; 1469467 CN;  
SQ  
Query Match 100.0%; Score 3806; DB 14; Length 510;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MEDTKESNVKTFCSKNILALIGFSSIIAVIALAVGLTQNKALPENVKYGIYLDAGSSHT 60



D	b	1	MEDRKESVWKEGCKNIIAIGFSSIIIAVALAIVGLTONKALPENKYGIVDAGSSHT	60
Q	y	1	MEDRKESVWKEGCKNIIAIGFSSIIIAVALAIVGLTONKALPENKYGIVDAGSSHT	60
D	b	61	SLYITKMPAEKENDTVGVHGYEECRVKGPGISKEFVQVNEIGIYLTDCMERAREVIRSQ	120
Q	y	61	SLYITKMPAEKENDTVGVHGYEECRVKGPGISKEFVQVNEIGIYLTDCMERAREVIRSQ	120
D	b	121	HOETPYVIGAGNRLLRMESEELADRYLDVYVESLSNPPDROGAAITIGQEEGATGWT	180
Q	y	121	HOETPYVIGAGNRLLRMESEELADRYLDVYVESLSNPPDROGAAITIGQEEGATGWT	180
D	b	181	TINILLGKFSQKTFMESIVPEYETNNQETFGALDYGASTQYTFVPOQNTIESPDNALQFR	240
Q	y	181	TINILLGKFSQKTFMESIVPEYETNNQETFGALDYGASTQYTFVPOQNTIESPDNALQFR	240
D	b	241	LYGKDYNVYTHSFLCYGKDQALMOKLAKIDIOVASNEILRDCPHFGYKKVYVNSDLYKTP	300
Q	y	241	LYGKDYNVYTHSFLCYGKDQALMOKLAKIDIOVASNEILRDCPHFGYKKVYVNSDLYKTP	300
D	b	301	CTKRPEMTLPPOQEIOIGIGNYOQCHOSIIELEFNTSCPYSOCAFNGNITPLPLOGDGFAR	360
Q	y	301	CTKRPEMTLPPOQEIOIGIGNYOQCHOSIIELEFNTSCPYSOCAFNGNITPLPLOGDGFAR	360
D	b	361	SAFYFVVKFELNLTSEKVSQOEKVTBMMKRFCAQOPWEIKTISYAGVKEKYLSSEYCFSGTYIL	420
Q	y	361	SAFYFVVKFELNLTSEKVSQOEKVTBMMKRFCAQOPWEIKTISYAGVKEKYLSSEYCFSGTYIL	420
D	b	421	SLLOGVHFTDSEEHIFHGKIOGSDAGWITGMLNTJNMPAEQLOPLSLSTHYVEL	480
Q	y	421	SLLOGVHFTDSEEHIFHGKIOGSDAGWITGMLNTJNMPAEQLOPLSLSTHYVEL	480
D	b	481	MYEFLVLEFVAIIGLLIFHRKPSYEMKDMV	510

RESULT	5	STANDARD:	PRT:	464	AA.
ID	PCT-US99-22955-27				
XX					
XX	xxxxxx				
AC					
DT					
XX					
XX					
DE	Sequence 27, Application PC/TUS9922955				
XX					
CC	Sequence 27, Application PC/TUS9922955				
CC	GENERAL INFORMATION:				
CC	APPLICANT: Maliszewski, Charles R.				
CC	APPLICANT: Gayle III, Richard B.				
CC	APPLICANT: Price, Virginia L.				
CC	APPLICANT: Gimpel, Steven D.				
CC	APPLICANT: Immunex Corporation				
CC	TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment				
CC	FILE REFERENCE: 2879-WO				
CC	CURRENT APPLICATION NUMBER: PCT/US99/22955				
CC	CURRENT FILING DATE: 1999-10-13				
CC	EARLIER APPLICATION NUMBER: US 60/104,585				
CC	EARLIER FILING DATE: 1998-10-16				
CC	EARLIER APPLICATION NUMBER: US 60/107,466				
CC	EARLIER FILING DATE: 1998-11-06				
CC	EARLIER APPLICATION NUMBER: US 60/149,010				
CC	EARLIER FILING DATE: 1999-06-13				
CC	NUMBER OF SEQ ID NOS: 31				
CC	SOFTWARE: PatentIn Ver. 2.0				
CC	SEQ ID NO 27				
CC	LENGTH: 464				
CC	TYPE: PRT				
CC	ORGANISM: Artificial Sequence				
CC	FEATURE:				
CC	OTHER INFORMATION: Description of Artificial Sequence: Fusion				
CC	OTHER INFORMATION: construct of human CD39				

SQ SEQUENCE 464 AA; 52754 MW; 1182698 CN;  
Query Match 86.5%; Score 3292; DB 1; Length 464;  
Best Local Similarity 97.6%; Pred. No. 0.00e+00;  
Matches 443; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Db 11 LSCIALSLATVNSATONKALPENKYGIVLDAGSSHTSLYIKWPAEKENDTGVAHQVE 70  
: : : : :  
23 FSSIAVIALALVGLTQNKALPENKYGIVLDAGSSHTSLYIKWPAEKENDTGVAHQVE 82  
Db 71 ECRKVGPGISKRVQVQVNEIGIYITDCMERAREVIPSQHOETPVYLGATAGMRLMSE 130  
83 ECRKVGPGISKRVQVQVNEIGIYITDCMERAREVIPSQHOETPVYLGATAGMRLMSE 142  
Db 131 ELADRVLDIVERSLSNYPDFOGARLITGOEGAGWITINLLGFSQKTRMFSTVPE 190  
143 ELADRVLDIVERSLSNYPDFOGARLITGOEGAGWITINLLGFSQKTRMFSTVPE 202  
Db 191 TNNQETFGALDGGASTOVTFVPONOTIESPDNALQFLRYGKDYNYTHSFLCYGKQAL 250  
203 TNNQETFGALDGGASTOVTFVPONOTIESPDNALQFLRYGKDYNYTHSFLCYGKQAL 262  
Db 251 WQKLAKDIOVASNEILRDPCHFPGYKRVVNSDLYKTPCTKRREMTLPQOFEIOGIGNY 310  
263 WQKLAKDIOVASNEILRDPCHFPGYKRVVNSDLYKTPCTKRREMTLPQOFEIOGIGNY 322  
Db 311 QOCHOSILELFTSYCPYSOCAPNGIFLPLQDGFAGFAFYVMKFLNLTSEKVSQEKY 370  
323 QOCHOSILELFTSYCPYSOCAPNGIFLPLQDGFAGFAFYVMKFLNLTSEKVSQEKY 382  
Db 371 TEMMKRFGAQPMEIKTSYAGVKEKYLSEYCSGTYYLSLLQGYHFTADSWEHIFICK 430  
383 TEMMKRFGAQPMEIKTSYAGVKEKYLSEYCSGTYYLSLLQGYHFTADSWEHIFICK 442  
Db 431 IQGSDAGWTLGMLNLTNMPAEOPLSTPLSHST 464  
443 IQGSDAGWTLGMLNLTNMPAEOPLSTPLSHST 476

RESULT 6 STANDARD; PRT; 439 AA.  
ID US-09-374-586-2  
AC xxxxxx  
XX  
DT  
XX  
XX  
DE Sequence 2, Application US/09374586  
XX  
CC Sequence 2, Application US/09374586  
CC GENERAL INFORMATION:  
CC APPLICANT: Pinsky, David J.  
CC TITLE OF INVENTION: CD39/ECTO-ADPASE AS A TREATMENT FOR THROMBOTIC AND  
CC FILE REFERENCE: 0575/59167  
CC CURRENT APPLICATION NUMBER: US/09/374,586  
CC CURRENT FILING DATE: 1999-08-13  
CC NUMBER OF SEQ ID NOS: 2  
CC SOFTWARE: Patentln Ver. 2.1  
CC SEQ ID NO 2  
CC LENGTH: 439  
CC TYPE: PRT  
CC ORGANISM: Homo sapiens  
SQ SEQUENCE 439 AA; 50038 MW; 1060957 CN;

Query Match 86.0%; Score 3275; DB 25; Length 439;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TONKALPENKYGIVLDAGSSHTSLYIKWPAEKENDTGVAHQVEECRKVGPGISKRVQK 60  
38 TONKALPENKYGIVLDAGSSHTSLYIKWPAEKENDTGVAHQVEECRKVGPGISKRVQK 97  
Db 61 VNEIGIYITDCMERAREVIPSQHOETPVYLGATAGMRLMSEBELADRVLDIVERSLS 120

OY 98 VNEIGIYITDCMERAREVIPSQHOETPVYLGATAGMRLMSEBELADRVLDIVERSLS 157  
Db 121 NYPDFOGARLITGOEGAGWITINLLGFSQKTRMFSTVPEETNNOETFGALDGG 180  
OY 158 NYPDFOGARLITGOEGAGWITINLLGFSQKTRMFSTVPEETNNOETFGALDGG 217  
Db 181 STOVTFVPONOTIESPDNALQFLRYGKDYNYTHSFLCYGKQDALMOKLADIOVASNEI 240  
OY 218 STOVTFVPONOTIESPDNALQFLRYGKDYNYTHSFLCYGKQDALMOKLADIOVASNEI 277  
Db 241 LNDPCFHPGKRVVNSDLYKTPCTKRREMTLPQOFEIOGIGNYQOCHOSILELFTSY 300  
OY 278 LNDPCFHPGKRVVNSDLYKTPCTKRREMTLPQOFEIOGIGNYQOCHOSILELFTSY 337  
Db 301 CPYSOCAPNGIFLPLQDGFAGFAFYVMKFLNLTSEKVSQEKYTEMMKRFCAQPMEEI 360  
OY 338 CPYSOCAPNGIFLPLQDGFAGFAFYVMKFLNLTSEKVSQEKYTEMMKRFCAQPMEEI 397  
Db 361 KTSYAGVKEKYLSEYCSGTYYLSLLQGYHFTADSWEHIFICKIQGSDAGWTLGYMLN 420  
OY 398 KTSYAGVKEKYLSEYCSGTYYLSLLQGYHFTADSWEHIFICKIQGSDAGWTLGYMLN 457  
Db 421 LTNMTPAEOPLSTPLSHST 439  
OY 458 LTNMTPAEOPLSTPLSHST 476

RESULT 7 STANDARD; PRT; 454 AA.  
ID PCT-US99-22955-6  
AC xxxxxx  
XX  
DT  
XX  
XX  
DE Sequence 6, Application PC/TUS9922955  
XX  
CC Sequence 6, Application PC/TUS9922955  
CC GENERAL INFORMATION:  
CC APPLICANT: Maliszewski, Charles R.  
CC APPLICANT: Gayle III, Richard B.  
CC APPLICANT: Gayle, Virginia L.  
CC APPLICANT: Gimpe, Steven D.  
CC APPLICANT: Immunex Corporation  
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment  
CC FILE REFERENCE: 2879-WO  
CC CURRENT APPLICATION NUMBER: PCT/US99/22955  
CC CURRENT FILING DATE: 1999-10-13  
CC EARLIER APPLICATION NUMBER: US 60/104,585  
CC EARLIER FILING DATE: 1998-10-16  
CC EARLIER APPLICATION NUMBER: US 60/107,466  
CC EARLIER FILING DATE: 1998-11-06  
CC EARLIER APPLICATION NUMBER: US 60/149,010  
CC NUMBER OF SEQ ID NOS: 31  
CC SOFTWARE: Patentln Ver. 2.0  
CC SEQ ID NO 6  
CC LENGTH: 454  
CC TYPE: PRT  
CC ORGANISM: Artificial Sequence  
SQ SEQUENCE 454 AA; 51543 MW; 1133800 CN;

Query Match 86.0%; Score 3275; DB 1; Length 454;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 16 TONKALPENKYGIVLDAGSSHTSLYIKWPAEKENDTGVAHQVEECRKVGPGISKRVQK 75  
OY 38 TONKALPENKYGIVLDAGSSHTSLYIKWPAEKENDTGVAHQVEECRKVGPGISKRVQK 97  
Db 76 VNEIGIYITDCMERAREVIPSQHOETPVYLGATAGMRLMSEBELADRVLDIVERSLS 135  
OY 98 VNEIGIYITDCMERAREVIPSQHOETPVYLGATAGMRLMSEBELADRVLDIVERSLS 157

Db 136 NYPPDFOGARITIGQEGGANGWITINLLGKFSQKTRWFSTVPEETNNQETFGALDLGGA 195  
|  
QY 158 NYPPDFOGARITIGQEGGANGWITINLLGKFSQKTRWFSTVPEETNNQETFGALDLGGA 217  
|  
Db 196 STQVTFVPOQNOTIESPNNALQFRLYKGDYNYTHSFCLCYGKDQALMOKLAKDIOVASNEI 255  
|  
QY 218 STQVTFVPOQNOTIESPNNALQFRLYKGDYNYTHSFCLCYGKDQALMOKLAKDIOVASNEI 277  
|  
Db 256 LRDPCHFPGYKRVVNSDLYKTPCTKRFEMLTPROQFEIIGIGNYQOCHOSILELFNTSY 315  
|  
QY 278 LRDPCHFPGYKRVVNSDLYKTPCTKRFEMLTPROQFEIIGIGNYQOCHOSILELFNTSY 337  
|  
Db 316 CPYSQCAFNGIFLPLPDGDFGAFSAFYFVKFLLNLTSKVSQEKVTEMMKFCQAPWEEI 375  
|  
QY 338 CPYSQCAFNGIFLPLPDGDFGAFSAFYFVKFLLNLTSKVSQEKVTEMMKFCQAPWEEI 397  
|  
Db 376 KTSYAGVKEKYLSEYCSFGTYILSLLOGYHFTADSMHIFHFGKIGSDAGWTLGYMLN 435  
|  
QY 398 KTSYAGVKEKYLSEYCSFGTYILSLLOGYHFTADSMHIFHFGKIGSDAGWTLGYMLN 457  
|  
Db 436 LTNMIPAEOPLSTPLSHST 454  
|  
QY 458 LTNMIPAEOPLSTPLSHST 476

RESULT 8  
ID PCT-US99-22955-30 STANDARD; PRT: 463 AA.

AC xxxxxx

Sequence 30, Application PC/TUS9922955

CC GENERAL INFORMATION:

CC APPLICANT: Maliszewski, Charles R.

CC APPLICANT: Gayle III, Richard B.

CC APPLICANT: Price, Virginia L.

CC APPLICANT: Gimpel, Steven D.

CC APPLICANT: Immunex Corporation

CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment

CC FILE REFERENCE: 2879-WO

CC CURRENT APPLICATION NUMBER: PCT/US99/22955

CC EARLIER FILING DATE: 1999-10-13

CC EARLIER APPLICATION NUMBER: US 60/104,585

CC EARLIER FILING DATE: 1998-10-16

CC EARLIER APPLICATION NUMBER: US 60/107,466

CC EARLIER FILING DATE: 1998-11-06

CC EARLIER APPLICATION NUMBER: US 60/149,010

CC EARLIER FILING DATE: 1999-08-13

CC NUMBER OF SEQ ID NOS: 31

CC SOFTWARE: Patentln Ver. 2.0

CC SEQ ID NO 30

CC LENGTH: 463

CC TYPE: PRT

CC ORGANISM: Artificial Sequence

CC FEATURE:

CC OTHER INFORMATION: Description of Artificial Sequence: Fusion

CC OTHER INFORMATION: construct of human CD39

CC SEQUENCE 463 AA: 52621 MW: 1178782 CN:

Query Match 86.0%; Score 3275; DB 1; Length 463;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;

Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 25 TONKALPENYKYGIVLDAGSSHTSLYIKWPAEKENDTGYYHGYVEBCRVKPGISKRYOK 84  
|  
QY 38 TONKALPENYKYGIVLDAGSSHTSLYIKWPAEKENDTGYYHGYVEBCRVKPGISKRYOK 97  
|  
Db 85 VNEIGIYLTDCMERAREVYIPRSQHOETPPVYLGAATAGMRLLRMESEBELADRYLDVVERSL 144  
|

QY 98 VNEIGIYLTDCMERAREVYIPRSQHOETPPVYLGAATAGMRLLRMESEBELADRYLDVVERSL 157  
|  
Db 145 NYPPDFOGARITIGQEGGANGWITINLLGKFSQKTRWFSTVPEETNNQETFGALDLGGA 204  
|  
QY 158 NYPPDFOGARITIGQEGGANGWITINLLGKFSQKTRWFSTVPEETNNQETFGALDLGGA 217  
|  
Db 205 STQVTFVPOQNOTIESPNNALQFRLYKGDYNYTHSFCLCYGKDQALMOKLAKDIOVASNEI 264  
|  
QY 218 STQVTFVPOQNOTIESPNNALQFRLYKGDYNYTHSFCLCYGKDQALMOKLAKDIOVASNEI 277  
|  
Db 265 LRDPCHFPGYKRVVNSDLYKTPCTKRFEMLTPROQFEIIGIGNYQOCHOSILELFNTSY 324  
|  
QY 278 LRDPCHFPGYKRVVNSDLYKTPCTKRFEMLTPROQFEIIGIGNYQOCHOSILELFNTSY 337  
|  
Db 325 CPYSQCAFNGIFLPLPDGDFGAFSAFYFVKFLLNLTSKVSQEKVTEMMKFCQAPWEEI 384  
|  
QY 338 CPYSQCAFNGIFLPLPDGDFGAFSAFYFVKFLLNLTSKVSQEKVTEMMKFCQAPWEEI 397  
|  
Db 385 KTSYAGVKEKYLSEYCSFGTYILSLLOGYHFTADSMHIFHFGKIGSDAGWTLGYMLN 444  
|  
QY 398 KTSYAGVKEKYLSEYCSFGTYILSLLOGYHFTADSMHIFHFGKIGSDAGWTLGYMLN 457  
|  
Db 445 LTNMIPAEOPLSTPLSHST 463  
|  
QY 458 LTNMIPAEOPLSTPLSHST 476

RESULT 9  
ID PCT-US99-22955-29 STANDARD; PRT: 473 AA.

AC xxxxxx

Sequence 29, Application PC/TUS9922955

CC GENERAL INFORMATION:

CC APPLICANT: Maliszewski, Charles R.

CC APPLICANT: Gayle III, Richard B.

CC APPLICANT: Price, Virginia L.

CC APPLICANT: Gimpel, Steven D.

CC APPLICANT: Immunex Corporation

CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment

CC FILE REFERENCE: 2879-WO

CC CURRENT APPLICATION NUMBER: PCT/US99/22955

CC EARLIER FILING DATE: 1999-10-13

CC EARLIER APPLICATION NUMBER: US 60/104,585

CC EARLIER FILING DATE: 1998-10-16

CC EARLIER APPLICATION NUMBER: US 60/107,466

CC EARLIER FILING DATE: 1998-11-06

CC EARLIER APPLICATION NUMBER: US 60/149,010

CC EARLIER FILING DATE: 1999-08-13

CC NUMBER OF SEQ ID NOS: 31

CC SOFTWARE: Patentln Ver. 2.0

CC SEQ ID NO 29

CC LENGTH: 473

CC TYPE: PRT

CC ORGANISM: Artificial Sequence

CC FEATURE:

CC OTHER INFORMATION: Description of Artificial Sequence: Fusion

CC OTHER INFORMATION: construct of human CD39

CC SEQUENCE 473 AA: 53745 MW: 1229492 CN:

Query Match 86.0%; Score 3275; DB 1; Length 473;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;

Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 35 TONKALPENYKYGIVLDAGSSHTSLYIKWPAEKENDTGYYHGYVEBCRVKPGISKRYOK 94  
|  
QY 38 TONKALPENYKYGIVLDAGSSHTSLYIKWPAEKENDTGYYHGYVEBCRVKPGISKRYOK 97  
|  
Db 95 VNEIGIYLTDCMERAREVYIPRSQHOETPPVYLGAATAGMRLLRMESEBELADRYLDVVERSL 154  
|

QY 98 VNEIGIYLTDCMERAREVYIPRSOHOETPYLGATAGMRLRMESEBELADRVLDVVERSL 157  
155 NYPDFOGARIIITGOEAGYGMWITINYLKFSQKTRMFSTIVPYETNNQETFGALDIGA 214  
QY 158 NYPDFOGARIIITGOEAGYGMWITINYLKFSQKTRMFSTIVPYETNNQETFGALDIGA 217  
Db 215 STQVTFVPONOTIESPDNALQFRLYKGDYNYVTHSFLLCYGKQALMOKLADIVASNEI 274  
QY 218 STQVTFVPONOTIESPDNALQFRLYKGDYNYVTHSFLLCYGKQALMOKLADIVASNEI 277  
Db 275 LRDPCFHPGKRVVNSDLKTPCTKREEMTLPFQOFIEIGIGNYQOCHQSILELFTSY 334  
QY 278 LRDPCFHPGKRVVNSDLKTPCTKREEMTLPFQOFIEIGIGNYQOCHQSILELFTSY 337  
Db 335 CPYSOCAFNGIFLPLQGDGFAFSAFYVMKFLNTSEKVSQEKYTEMMKFCQAOPMEI 394  
QY 338 CPYSOCAFNGIFLPLQGDGFAFSAFYVMKFLNTSEKVSQEKYTEMMKFCQAOPMEI 397  
Db 395 KTSYAGVKEKYLSEYCFSGTYILSLLOGYHFTADSWEHIFHIGKIQSDAGMTLGMYLN 454  
QY 398 KTSYAGVKEKYLSEYCFSGTYILSLLOGYHFTADSWEHIFHIGKIQSDAGMTLGMYLN 457  
Db 455 LTNMIPAEQPLSTPLSHST 473  
QY 458 LTNMIPAEQPLSTPLSHST 476

RESULT 10  
ID PCT-US99-22955-28 STANDARD: PRT: 474 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE Sequence 28, Application PC/TUS9922955  
XX  
CC Sequence 28, Application PC/TUS9922955  
CC GENERAL INFORMATION:  
CC APPLICANT: Maliszewski, Charles R.  
CC APPLICANT: Gayle III, Richard B.  
CC APPLICANT: Gayle, Virginia L.  
CC APPLICANT: Price, Steven D.  
CC APPLICANT: Gimpel, Steven D.  
CC APPLICANT: Immunex Corporation  
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment  
CC FILE REFERENCE: 2879-WO  
CC CURRENT APPLICATION NUMBER: PCT/US99/22955  
CC CURRENT FILING DATE: 1999-10-13  
CC EARLIER APPLICATION NUMBER: US 60/104,585  
CC EARLIER FILING DATE: 1998-10-16  
CC EARLIER APPLICATION NUMBER: US 60/107,466  
CC EARLIER FILING DATE: 1998-11-06  
CC EARLIER APPLICATION NUMBER: US 60/149,010  
CC EARLIER FILING DATE: 1999-08-13  
CC NUMBER OF SEQ ID NOS: 31  
CC SOFTWARE: Patentln Ver. 2.0  
CC SEQ ID NO 28  
CC LENGTH: 474  
CC TYPE: PRT  
CC ORGANISM: Artificial Sequence  
CC FEATURE:  
CC OTHER INFORMATION: Description of Artificial Sequence: Fusion  
CC OTHER INFORMATION: construct of human CD39  
CC SEQUENCE 474 AA: 53816 MW: 1234400 CN:

Query Match 86.0%; Score 3275; DB 1; Length 474;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 36 TONKALPENVKYGYVLADGSSHTSLYIKWPAEKENDTGVAHVQVECRVKGPGISKPFVOK 95  
QY 38 TONKALPENVKYGYVLADGSSHTSLYIKWPAEKENDTGVAHVQVECRVKGPGISKPFVOK 97

Db 96 VNEIGIYLTDCMERAREVYIPRSOHOETPYLGATAGMRLRMESEBELADRVLDVVERSL 155  
QY 98 VNEIGIYLTDCMERAREVYIPRSOHOETPYLGATAGMRLRMESEBELADRVLDVVERSL 157  
Db 156 NYPDFOGARIIITGOEAGYGMWITINYLKFSQKTRMFSTIVPYETNNQETFGALDIGA 215  
QY 158 NYPDFOGARIIITGOEAGYGMWITINYLKFSQKTRMFSTIVPYETNNQETFGALDIGA 217  
Db 216 STQVTFVPONOTIESPDNALQFRLYKGDYNYVTHSFLLCYGKQALMOKLADIVASNEI 275  
QY 218 STQVTFVPONOTIESPDNALQFRLYKGDYNYVTHSFLLCYGKQALMOKLADIVASNEI 277  
Db 276 LRDPCFHPGKRVVNSDLKTPCTKREEMTLPFQOFIEIGIGNYQOCHQSILELFTSY 335  
QY 278 LRDPCFHPGKRVVNSDLKTPCTKREEMTLPFQOFIEIGIGNYQOCHQSILELFTSY 337  
Db 336 CPYSOCAFNGIFLPLQGDGFAFSAFYVMKFLNTSEKVSQEKYTEMMKFCQAOPMEI 395  
QY 338 CPYSOCAFNGIFLPLQGDGFAFSAFYVMKFLNTSEKVSQEKYTEMMKFCQAOPMEI 397  
Db 396 KTSYAGVKEKYLSEYCFSGTYILSLLOGYHFTADSWEHIFHIGKIQSDAGMTLGMYLN 455  
QY 398 KTSYAGVKEKYLSEYCFSGTYILSLLOGYHFTADSWEHIFHIGKIQSDAGMTLGMYLN 457  
Db 456 LTNMIPAEQPLSTPLSHST 474  
QY 458 LTNMIPAEQPLSTPLSHST 476

RESULT 11  
ID PCT-US99-22955-3 STANDARD: PRT: 476 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE Sequence 3, Application PC/TUS9922955  
XX  
CC Sequence 3, Application PC/TUS9922955  
CC GENERAL INFORMATION:  
CC APPLICANT: Maliszewski, Charles R.  
CC APPLICANT: Gayle III, Richard B.  
CC APPLICANT: Price, Steven D.  
CC APPLICANT: Gimpel, Steven D.  
CC APPLICANT: Immunex Corporation  
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment  
CC FILE REFERENCE: 2879-WO  
CC CURRENT APPLICATION NUMBER: PCT/US99/22955  
CC CURRENT FILING DATE: 1999-10-13  
CC EARLIER APPLICATION NUMBER: US 60/104,585  
CC EARLIER FILING DATE: 1998-10-16  
CC EARLIER APPLICATION NUMBER: US 60/107,466  
CC EARLIER FILING DATE: 1998-11-06  
CC EARLIER APPLICATION NUMBER: US 60/149,010  
CC EARLIER FILING DATE: 1999-08-13  
CC NUMBER OF SEQ ID NOS: 31  
CC SOFTWARE: Patentln Ver. 2.0  
CC SEQ ID NO 3  
CC LENGTH: 476  
CC TYPE: PRT  
CC ORGANISM: Artificial Sequence  
CC FEATURE:  
CC OTHER INFORMATION: Description of Artificial Sequence: Fusion  
CC OTHER INFORMATION: construct of human CD39  
CC SEQUENCE 476 AA: 54177 MW: 1244820 CN:

Query Match 86.0%; Score 3275; DB 1; Length 476;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 38 TONKALPENVKYGYVLADGSSHTSLYIKWPAEKENDTGVAHVQVECRVKGPGISKPFVOK 97  
QY 38 TONKALPENVKYGYVLADGSSHTSLYIKWPAEKENDTGVAHVQVECRVKGPGISKPFVOK 97

Db 98 VNEIGIYLTDCMERAREVIRRSQHETPVYLGATAGMRLLRMESEELADRYLDVVERSLS 157  
|  
Qy 98 VNEIGIYLTDCMERAREVIRRSQHETPVYLGATAGMRLLRMESEELADRYLDVVERSLS 157  
|  
Db 158 NYPEDFOGARITIGOEAGAYGWITINYLKFSQKTRMFSTIVPEYETNNQETFGALDYGGA 217  
|  
Qy 158 NYPEDFOGARITIGOEAGAYGWITINYLKFSQKTRMFSTIVPEYETNNQETFGALDYGGA 217  
|  
Db 218 STQVTFVPOQNTIESPNDALQFRLYGKDYNYTHSFCLCYGKDQALMOKLAKDIOVASNEI 277  
|  
Qy 218 STQVTFVPOQNTIESPNDALQFRLYGKDYNYTHSFCLCYGKDQALMOKLAKDIOVASNEI 277  
|  
Db 278 LRDCCFHFGKRYKVVNSDLKTKPCTKRFEMLTPPOQFEIIGIGNYOQCHOSILELFNTSY 337  
|  
Qy 278 LRDCCFHFGKRYKVVNSDLKTKPCTKRFEMLTPPOQFEIIGIGNYOQCHOSILELFNTSY 337  
|  
Db 338 CPYSQCAFNGIFLPLDGDGAFSAFYVVKFNLJTSEKVSQEKVTBMKRFCAQPMEEI 397  
|  
Qy 338 CPYSQCAFNGIFLPLDGDGAFSAFYVVKFNLJTSEKVSQEKVTBMKRFCAQPMEEI 397  
|  
Db 398 KTSYAGVKEKYLSEYCSGYIISLLQGYHFTADSWEHIFHFGKIQGSDAGWTLYGMLN 457  
|  
Qy 398 KTSYAGVKEKYLSEYCSGYIISLLQGYHFTADSWEHIFHFGKIQGSDAGWTLYGMLN 457  
|  
Db 458 LTNMIRPAEQPLSTPLSHST 476  
|  
Qy 458 LTNMIRPAEQPLSTPLSHST 476  
|

RESULT 12  
ID PCT-US99-22955-8 STANDARD: PRT: 478 AA.

AC xxxxxx

Sequence 8, Application PC/TUS9922955

CC Sequence 8, Application PC/TUS9922955  
CC GENERAL INFORMATION:  
CC APPLICANT: Maliszewski, Charles R.  
CC APPLICANT: Gayle III, Richard B.  
CC APPLICANT: Price, Virginia L.  
CC APPLICANT: Gimpel, Steven D.  
CC APPLICANT: Immunex Corporation  
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment  
CC FILE REFERENCE: 2879-WO  
CC CURRENT APPLICATION NUMBER: PCT/US99/22955  
CC EARLIER FILING DATE: 1999-10-13  
CC EARLIER APPLICATION NUMBER: US 60/104,585  
CC EARLIER FILING DATE: 1998-10-16  
CC EARLIER APPLICATION NUMBER: US 60/107,466  
CC EARLIER FILING DATE: 1998-11-06  
CC EARLIER APPLICATION NUMBER: US 60/149,010  
CC EARLIER FILING DATE: 1999-08-13  
CC NUMBER OF SEQ ID NOS: 31  
CC SOFTWARE: Patentln Ver. 2.0  
CC SEQ ID NO 8  
CC LENGTH: 478  
CC TYPE: PRT  
CC ORGANISM: Artificial Sequence  
CC SEQUENCE 478 AA; 54188 MW; 1255693 CN;

Query Match 86.0%; Score 3275; DB 1; Length 478;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 40 TONKALPENVKYGIYLDAGSSHTSLYTKWPAEKENDTGVAHVQECRVKPGISKVQK 99  
|  
Qy 38 TONKALPENVKYGIYLDAGSSHTSLYTKWPAEKENDTGVAHVQECRVKPGISKVQK 97  
|  
Db 100 VNEIGIYLTDCMERAREVIRRSQHETPVYLGATAGMRLLRMESEELADRYLDVVERSLS 159  
|

Qy 98 VNEIGIYLTDCMERAREVIRRSQHETPVYLGATAGMRLLRMESEELADRYLDVVERSLS 157  
|  
Db 160 NYPEDFOGARITIGOEAGAYGWITINYLKFSQKTRMFSTIVPEYETNNQETFGALDYGGA 219  
|  
Qy 158 NYPEDFOGARITIGOEAGAYGWITINYLKFSQKTRMFSTIVPEYETNNQETFGALDYGGA 217  
|  
Db 220 STQVTFVPOQNTIESPNDALQFRLYGKDYNYTHSFCLCYGKDQALMOKLAKDIOVASNEI 279  
|  
Qy 218 STQVTFVPOQNTIESPNDALQFRLYGKDYNYTHSFCLCYGKDQALMOKLAKDIOVASNEI 277  
|  
Db 280 LRDCCFHFGKRYKVVNSDLKTKPCTKRFEMLTPPOQFEIIGIGNYOQCHOSILELFNTSY 339  
|  
Qy 278 LRDCCFHFGKRYKVVNSDLKTKPCTKRFEMLTPPOQFEIIGIGNYOQCHOSILELFNTSY 337  
|  
Db 340 CPYSQCAFNGIFLPLDGDGAFSAFYVVKFNLJTSEKVSQEKVTBMKRFCAQPMEEI 399  
|  
Qy 338 CPYSQCAFNGIFLPLDGDGAFSAFYVVKFNLJTSEKVSQEKVTBMKRFCAQPMEEI 397  
|  
Db 400 KTSYAGVKEKYLSEYCSGYIISLLQGYHFTADSWEHIFHFGKIQGSDAGWTLYGMLN 459  
|  
Qy 398 KTSYAGVKEKYLSEYCSGYIISLLQGYHFTADSWEHIFHFGKIQGSDAGWTLYGMLN 457  
|  
Db 460 LTNMIRPAEQPLSTPLSHST 478  
|  
Qy 458 LTNMIRPAEQPLSTPLSHST 476  
|

RESULT 13  
ID PCT-US99-22955-26 STANDARD: PRT: 487 AA.

AC xxxxxx

Sequence 26, Application PC/TUS9922955

CC Sequence 26, Application PC/TUS9922955  
CC GENERAL INFORMATION:  
CC APPLICANT: Maliszewski, Charles R.  
CC APPLICANT: Gayle III, Richard B.  
CC APPLICANT: Price, Virginia L.  
CC APPLICANT: Gimpel, Steven D.  
CC APPLICANT: Immunex Corporation  
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment  
CC FILE REFERENCE: 2879-WO  
CC CURRENT APPLICATION NUMBER: PCT/US99/22955  
CC EARLIER FILING DATE: 1999-10-13  
CC EARLIER APPLICATION NUMBER: US 60/104,585  
CC EARLIER FILING DATE: 1998-10-16  
CC EARLIER APPLICATION NUMBER: US 60/107,466  
CC EARLIER FILING DATE: 1998-11-06  
CC EARLIER APPLICATION NUMBER: US 60/149,010  
CC EARLIER FILING DATE: 1999-08-13  
CC NUMBER OF SEQ ID NOS: 31  
CC SOFTWARE: Patentln Ver. 2.0  
CC SEQ ID NO 26  
CC LENGTH: 487  
CC TYPE: PRT  
CC ORGANISM: Artificial Sequence  
CC SEQUENCE 487 AA; 55240 MW; 1301482 CN;

Query Match 86.0%; Score 3275; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

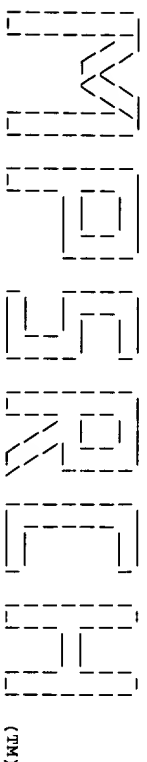
Db 49 TONKALPENVKYGIYLDAGSSHTSLYTKWPAEKENDTGVAHVQECRVKPGISKVQK 108  
|  
Qy 38 TONKALPENVKYGIYLDAGSSHTSLYTKWPAEKENDTGVAHVQECRVKPGISKVQK 97  
|  
Db 109 VNEIGIYLTDCMERAREVIRRSQHETPVYLGATAGMRLLRMESEELADRYLDVVERSLS 168  
|  
Qy 98 VNEIGIYLTDCMERAREVIRRSQHETPVYLGATAGMRLLRMESEELADRYLDVVERSLS 157  
|

Qy	49	YGVIVDASSHTSLIYTKWPAEKENDGVVHQVECRKAGPGISKFQYKNNEIGVITLDC	108
Db	109	MERAREV1PRSOHETPVYLICATAGMRLLRMESEBELADRVLDVVERLSLNPEDFOGARI	168
Oy	109	MERAREV1PRSOHETPVYLICATAGMRLLRMESEBELADRVLDVVERLSLNPEDFOGARI	168
Db	169	ITGOEEGAVGTTIIVYLIGKFSOKTRMFSIYPIETNNDEFGALDLCGASTOYTFVPOQ	228
Oy	169	ITGOEEGAVGTTIIVYLIGKFSOKTRMFSIYPIETNNDEFGALDLCGASTOYTFVPOQ	228
Db	229	TIESPDNALOFELRYKGDYNNVYTHSFLCYGKDQALMOKLAKDIQVANSNELLRDPCFHPGYK	288
Oy	229	TIESPDNALOFELRYKGDYNNVYTHSFLCYGKDQALMOKLAKDIQVANSNELLRDPCFHPGYK	288
Db	289	KVYVNSDYIKTPICIKREFMTLPFOOFIEQIGNTQQCHOSTILELFPNTSYCPYSQCAENG1	348
Oy	289	KVYVNSDYIKTPICIKREFMTLPFOOFIEQIGNTQQCHOSTILELFPNTSYCPYSQCAENG1	348
Db	349	FLPPLQGGPFGAFSAFYFMKFLNLTSEVSOEKTEMMKKRCQADPWEIITSVAGVEXY	408
Oy	349	FLPPLQGGPFGAFSAFYFMKFLNLTSEVSOEKTEMMKKRCQADPWEIITSVAGVEXY	408
Db	409	LSEYFSGSTYILSLLOGYHFTADSWEHIFHFGIKIQGSDAGWTGVLGNTLNMIIPAEOPL	468
Oy	409	LSEYFSGSTYILSLLOGYHFTADSWEHIFHFGIKIQGSDAGWTGVLGNTLNMIIPAEOPL	468
Db	469	STPLSHST 476	
Oy	469	STPLSHST 476	
RESULT	15		
ID	US-08-701-460-4	STANDARD;	PRT; 510 AA.
AC	xxxxxx		
DT			
DE			
xx			
xx			
Sequence 4, Application US/08701460			
CC	Sequence 4, Application US/08701460		
CC	GENERAL INFORMATION:		
CC	APPLICANT: MARCUS, AARON J.		
CC	APPLICANT: MALISZEWSKI, CHARLES R.		
CC	APPLICANT: GAYLE, RICHARD B.		
CC	TITLE OF INVENTION: METHODS OF REGULATING HEMOSTASIS AND		
CC	TITLE OF INVENTION: THROMBOSIS, AND COMPOUNDS USEFUL THEREFOR		
CC	NUMBER OF SEQUENCES: 16		
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESSEE: Immunex Corporation		
CC	STREET: 51 University Street		
CC	CITY: Seattle		
CC	STATE: WA		
CC	COUNTRY: USA		
CC	ZIP: 98101		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: Apple Power Macintosh 7200/90		
CC	OPERATING SYSTEM: Apple Operating System 7.5.3		
CC	SOFTWARE: Microsoft Word 6.0.1 for Power Macintosh		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/701,460		
CC	FILING DATE: 22 AUGUST 1996		
CC	CLASSIFICATION: 514		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: Perkins, Patricia Anne		
CC	REGISTRATION NUMBER: 34,693		
CC	REFERENCE/DOCKET NUMBER: 2808		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: 206-587-0430		
CC	INFORMATION FOR SEQ ID NO: 4:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 510 amino acids		









Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution Rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Fri May 5 08:41:38 2000; Maspar time 49.00 Seconds  
Tabular output not generated. 490.937 Million cell updates/sec

Title: >US-09-374-586-1  
Description: (1-510) from US09374586.pep  
Perfect Score: 3806  
Sequence: 1 MEDTKESNVKTFCSKNILAI.....VAITGLLPHKPSYWKDMV 510

Scoring table: PAM 150  
Gap 11  
Searched: 142080 seqs, 47172406 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir62  
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 51.082; Variance 106.884; scale 0.478  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	3806	100.0	510	2	I56242 lymphoid cell activat	0.00e+00
2	378	9.9	537	2	T16596 hypothetical protein	4.12e-44
3	360	9.5	630	2	S50463 hypothetical protein	7.21e-44
4	319	8.4	454	2	JC4616 apyrase (EC 3.6.1.5)	1.40e-33
5	293	7.7	1052	2	T04439 hypothetical protein	4.85e-29
6	291	7.6	455	2	S48859 nucleoside triphospha	1.08e-28
7	265	7.0	518	2	A40732 guanosine diphospha	3.06e-24
8	131	3.4	371	2	T05213 hypothetical protein	1.75e-03
9	125	3.3	597	2	T04954 hypothetical protein	1.06e-02
10	127	3.3	628	2	JN0620 nucleoside-triphospha	5.84e-03
11	117	3.1	528	2	UDP-glucuronosyltrans	1.07e-01
12	116	3.0	107	2	G69988 thioredoxin H1 homo	1.42e-01
13	113	3.0	919	2	S57786 hypothetical protein	3.29e-01
14	114	3.0	974	1	EXLNDP H+-transporting ATPas	2.49e-01
15	111	2.9	369	2	A72453 conserved hypothetical	5.70e-01
16	110	2.9	392	2	A72453 conserved hypothetical	5.70e-01
17	109	2.9	586	2	T12886 conserved hypothetical	9.81e-01
18	106	2.8	204	2	C58930 NADH dehydrogenase (u	2.19e+00
19	107	2.8	276	2	S13585 most protein precurs	1.68e+00
20	108	2.8	466	2	PC4296 nicotinic acetylcholi	1.28e+00
21	106	2.8	521	2	C28529 nicotinic acetylcholi	2.19e+00
22	106	2.8	573	2	F70860 probable cytochrome-c	2.19e+00
23	108	2.8	669	2	T05212 hypothetical protein	1.28e+00

24	104	2.7	44	2	S63501 apyrase (EC 3.6.1.5)	3.69e+00
25	101	2.7	222	2	C70020 conserved hypothetical	7.99e+00
26	104	2.7	296	2	A72239 conserved hypothetical	3.69e+00
27	102	2.7	450	2	JN0105 tubulin alpha-3 chain	6.19e+00
28	103	2.7	493	1	ACRYB1 nicotinic acetylcholi	4.79e+00
29	102	2.7	494	2	T03774 probable histidine-t	6.19e+00
30	101	2.7	501	2	S04607 nicotinic acetylcholi	7.99e+00
31	101	2.7	505	2	S07227 nicotinic acetylcholi	7.99e+00
32	102	2.7	530	2	A48633 glucuronosyltransfera	6.19e+00
33	104	2.7	544	2	T05211 hypothetical protein	3.69e+00
34	101	2.7	557	2	S12359 nicotinic acetylcholi	7.99e+00
35	103	2.7	571	1	RNCW7H transcription initiat	4.79e+00
36	101	2.7	571	1	RNCW7T transcription initiat	7.99e+00
37	102	2.7	574	2	T05214 hypothetical protein	6.19e+00
38	102	2.7	575	2	S75962 NADH dehydrogenase (u	6.19e+00
39	103	2.7	634	2	S72725 guanosine-3',5'bis(di	4.79e+00
40	102	2.7	790	2	F70725 probable reja protein	6.19e+00
41	103	2.7	867	2	G69485 DNA-directed RNA poly	4.79e+00
42	102	2.7	923	2	S44226 periodic tryptophan p	6.19e+00
43	104	2.7	1680	2	T01367 hypothetical protein	3.69e+00
44	101	2.7	2672	2	A48126 translation activator	7.99e+00
45	100	2.6	461	2	S68137 NADH dehydrogenase (u	1.03e+01

ALIGNMENTS

ENTRY	1	I56242	#type complete	lymphoid cell activation antigen - human
TITLE				#formal_name Homo sapiens #common_name man
ORGANISM				02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
DATE				02-Jul-1996
ACCESSIONS				I56242
REFERENCE				I56242
#authors				Maliszewski, C.R.; Delespesse, G.J.; Schoenborn, M.A.; Armitage, R.J.; Fanslow, W.C.; Nakajima, T.; Baker, E.; Sutherland, G.R.; Polindexter, K.; Birks, C.
#journal				J. Immunol. (1994) 153:3574-3583
#title				The CD39 lymphoid cell activation antigen. Molecular cloning and structural characterization.
#cross-references				MUID:95015846
#accession				I56242
#status				preliminary; translated from GB/EMBL/DBJ
#molecule_type				mRNA
#residues				1-510 #label RES
#cross-references				GB:S7813; NID:9765255; PID:9765256
SUMMARY				#length 510 #molecular_weight 57964 #checksum 4387
Query Match				100.0%; Score 3806; DB 2; Length 510;
Best Local Similarity				100.0%; Pred. No. 0.00e+00;
Matches				510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	1	MEDTKESNVKTFCSKNILAI	IGFSSIIAVIALAVGLTQNKALPENKYGIVLDAGSSHT	60
Qy	1	MEDTKESNVKTFCSKNILAI	IGFSSIIAVIALAVGLTQNKALPENKYGIVLDAGSSHT	60
Db	61	SLYTYKPAEKENDGVHVECECRVKGPGISKFVQVNEIGIYLTDCMERAREVIPSQ	120	
Qy	61	SLYTYKPAEKENDGVHVECECRVKGPGISKFVQVNEIGIYLTDCMERAREVIPSQ	120	
Db	121	HOETPVYLGATAGKRLRLMESEELADRVLDVVERSLSNYPDFOGARITIGQEGAGWI	180	
Qy	121	HOETPVYLGATAGKRLRLMESEELADRVLDVVERSLSNYPDFOGARITIGQEGAGWI	180	
Db	181	TINTLAKFSQTRMFSIVPEYETNNOETFGALDAGASTQVTFPQNOTIESPNALOFR	240	
Qy	181	TINTLAKFSQTRMFSIVPEYETNNOETFGALDAGASTQVTFPQNOTIESPNALOFR	240	
Db	241	LYGKDYVYTHSFICYGQDALMOKLAKDIOVASNEILRDPCEHPGKXKVVNSDLYKTP	300	
Qy	241	LYGKDYVYTHSFICYGQDALMOKLAKDIOVASNEILRDPCEHPGKXKVVNSDLYKTP	300	
Db	301	CTKFEKTLPPQPEIFIGISGNYOCHOSILEFTWTSICPSCAFNGLFLPPLQGDGCAF	360	

QY	301	CKRREMLPPOQFELQIGNGQCHQHSILFELFNSTYCPYQCAFNGILFEPPLQDGC	360
Db	361	SAFYVPMKFLNLTSEKVSQEKYTEMMKFCAPQWEIKTSTAGYKEKYLSEYCSGYIL	420
QY	361	SAFYVPMKFLNLTSEKVSQEKYTEMMKFCAPQWEIKTSTAGYKEKYLSEYCSGYIL	420
Db	421	SLLLQGYHTADSMWHIFHIGIKIQSSDQGWTLGWLNTLNNIIPAQPLSTPLSHSTYVL	480
QY	421	SLLLQGYHTADSMWHIFHIGIKIQSSDQGWTLGWLNTLNNIIPAQPLSTPLSHSTYVL	480
Db	481	MYLFSVLTFTVAIIIGLIFHKRSTYMKMAY 510	
QY	481	MYLFSVLTFTVAIIIGLIFHKRSTYMKMAY 510	
RESULT	2		
ENTRY	116696	#type complete	
TITLE	hypothetical protein R07E4.4 - Caenorhabditis elegans		
ORGANISM	#formal_name Caenorhabditis elegans		
DATE	20-Sep-1999	#sequence_revision 20-Sep-1999	#text_change
ACCESSIONS	116696		
REFERENCE	118561		
#authors	Miller, N.		
#submission	submitted to the EMBL Data Library, October 1995		
#description	The sequence of C. elegans consmid R07E4.		
#accession	116696		
#status	preliminary; translated from GB/EMBL/DBJ		
#molecule_type	DNA		
#residues	1-557	#label MLT.	
#cross_references	EMBL:U09652; NID:g1049390; PID:g1049394;		
	PIDN:AAA80403.1; CESP:R07E4.4		
GENETICS	#experimental_source strain Bristol N2		
#gene	CESP:R07E4.4		
#introns	39/1; 67/1; 145/2; 192/1; 244/3; 317/3; 451/3; 488/2		
SUMMARY	#length 557	#molecular_weight 63056	#checksum 4267
Query Match	9.9%; Score 378; DB 2; Length 557;		
Best Local Similarity	23.9%; Pred. No. 4,12e-44;		
Matches	122; Conservative 134; Mismatches 217; Indels 37; Gaps 30;		
Db	12	VSAMTFEPVI-VFIIYVBAHPSPKYIADDERSGVICIDAGSTGRFLFYNMISTDSEL 70	
QY	18	LAILGFSSIIAIVALLAVGLTQNKALPEN-VK-VGIYDADGSSHTSLYIKWPREKENDI 75	
Db	71	IQIEGVYIDNKRPMKKISIPGLSTPECTRAQAQAEYIRPLMELAEKHIPERKPTYPVIFA 130	
QY	76	GVVHQQE-ECR-V-K-GPGISKFPQKQNEIGIYLTDCMERAREVIRPSOHOETPVYIGA 130	
Db	131	TAGMMLIDPEYVLIOQKAAVILKLNKRLPKTISMOVLEKHEIRILEGKEGIYSIAVNYA 190	
QY	131	TAGMMLIMESSELADR-VLDYVERSLSNY-PPD-FQG-ARITIGQEEGAYIGMTITNYL 165	
Db	191	LGRFNKATLDPFGTSPAHAKQITGVIMDQGSASAQIAFELPDTDFSSINVENINLGR 250	
QY	186	LGRFSQKTRM-FSIVPYETNNQDEFGALDIGASQVQYF-VPAQOTIES-P-DW-ALQFR 240	
Db	251	EDDSLKFKKLVYTPFLGIVGVEGIRFKYIHMILSKLQONGVIVIDDCCPLNHTYVLEN 310	
QY	241	LYGK--DENVYTHSFLCGKQDQALM--QKLAQDIQVANSNILDPCFHPQYKRVVAVSD 295	
Db	311	G-ENF-VERGGNGNMTCSEVKKILN-PESSEVQCA-EAAKCYFGAVPAPISPLNIEM 366	
QY	296	LYKTKCTKRFEMETLPPOQFELQIGNGQCHQHSILFELFNSTYCPYQCAFNGILFEPPLQDGC 355	
Db	367	YG-FSEWYSTHDVLDGQO-YDAENIAKKTTOOYCSKRMSTIOAESKOLYPRADERL 423	
QY	356	DGFAASAFYFWK-FLNLTSEKVSQEKYTEMMKFCAPQW--E-EIKTS-YAVVKKYIL 409	
Db	424	RTOCFKSMITISVLDGTS-VDKTHNKQVSTIAGQEVOMALCAMYIHHKRPFLDSSR 482	

QY	410	SEYCSGYTILSLAQGVHPADSMENHIFGIQIGSDAGCTGLCYMLNTLMPIADQPLS	469
Db	483	NLIVKETHSSSESLMAPFLFSLANFCFLV	512
QY	470	PFLSHSTYVFLWFLFSLVLTVALIGLLIF	499
RESULT	3		
ENTRY		S50463	#type complete
TITLE		hypothetical protein YER005w - yeast (Saccharomyces cerevisiae)	
ORGANISM		#formal_name Saccharomyces cerevisiae	
DATE		28-May-1993 #sequence_revision 24-Feb-1995 #text_change 21-Nov-1997	
ACCESSIONS		S50463	
REFERENCE		S50428	
#authors		Dietrich, F.S.	
#submission		submitted to the EMBL Data Library, December 1994	
#description		The sequence of S. cerevisiae cosmids 9537, 9581, 9495, 9867 and lambda clone 5898.	
#accession		S50463	
#molecule_type		DNA	
#residues		1-630	#label DIE
#cross-references		EMBL:U18778; NID:g603592; PID:g603597; MIPS:YER005w	
GENETICS			
#map_position	5R	length 630	#molecular_weight 71851
SUMMARY			#checksum 7317
Query Match		9.5%; Score 360; DB 2; Length 630;	
Best Local Similarity		28.4%; Pred. No. 7,21e-41;	
Matches	106;	Conservative 93; Mismatches 137; Indels 37; Gaps 28;	
Db	52	HOEKDWTFKLNPGLSFEKKRPQDAYKSHIKPLDFAKNIPESHWSSCPVFIQATAGMR	111
QY	78	VHQVECECRYK-GPGISKFEVQKVNET-GIYLDCKMERAREVIPSQHOETPVYLCAATAGMR	135
Db	112	LLRPDIQSSIIDGLCOGLKHP-AEFLVBDCAQIOVIDTGEIEGLGYGLNLYLGHND-	169
QY	136	ILRMSEEL-LADRYLADVERSLSNPF-DFQGA-RITTGDEGAYGITTNYLLGKRSQ	192
Db	170	---YN--P-EVSDHFTTFPFMDMGASPTQIAPRDSGEIARHRDITIFLRVNSGLQK	223
QY	193	TRWEIIVYEYNNQTFPFALDLGGASTOYTVTPQNQT-IESP-DN-ALQF-R-LYKQ-D-	245
Db	224	MDVEVSTMLGEGANQARRRYLAOLINTLPENT-ND-YENDDFSTR-NLNDPCAPRGSSTD	280
QY	246	VNVYHSHSLCGKQDALQMLAKIDQVASNILNDPCFHGKYKKVNAVVSLLYKTP-CTKR	304
Db	281	FEFKQTT-FHIASGNYEQCTKSTYPLLLKNMPCDDPCLFNGVAPRIIDFANDKFIGT	338
QY	305	FEMTLRPFQOFLEQIGNYQOCHOSTLELFNTSY-CPYSQCAFNFICFLPPLQ-G-D-FGAF	360
Db	339	SEYWTADVDFLGE--YNPDKFSLSREPCNSWMTQILANSQDGVNSIPENFLDAC	397
QY	361	SNFYIVNKLFLNLISEKYSQEKVETEMMKFCAQWEEI-----KTSYAGKXYLSBYCF	414
Db	398	KGNVNLILHSGF	410
QY	415	SGTYILSLDLOGY	427
RESULT	4		
ENTRY		JC4616	#type complete
TITLE		apsyrase (EC 3.6.1.5) precursor - potato	
ALTERNATE_NAMES		adenylpyrophosphatase; ATP-diphosphohydrolase	
ORGANISM		#formal_name Solanum tuberosum #common_name potato	
DATE		10-May-1996 #sequence_revision 19-Jul-1996 #text_change 28-May-1999	
ACCESSIONS		JC4616; PCA147	
REFERENCE		JC4616	
#authors		Handa, M.; Guidotti, G.	
#journal		Biochem. Biophys. Res. Commun. (1996) 218:916-923	
#title		Purification and cloning of a soluble ATP-diphosphohydrolase	

REFERENCE	215359	Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoerg, W.
#authors	5	Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoerg, W.
#submission	1	Submitted to the Protein Sequence Database, April 1998
#accession	T04439	
##molecule_type	DNA	
##residues	1-1052	##label BEV
##cross-references	EMBL:AL021667	
##experimental_source	cultivar Columbia; BAC clone T18B16	
GENETICS		
##map_position	4	
#introns	79/3; 118/3; 217/1; 295/3; 396/3; 531/2; 815/3	
#note	T18B16.150	
SUMMARY	#length 1052	#molecular_weight 116388 #checksum 4955
Query Match	7.7%; Score 293; DB 2; Length 1052;	
Best Local Similarity	27.3%; Pred. No. 4,85e-29;	
Matches	115; Conservative 120; Mismatches 135; Indels 52; Gaps 38;	
Db	578	PGPKLVNNRGLGKTAIKPLIQMAEKQIPKNAHNTTSLFVATAGVRRK--PAD-SSW 634
QY	89	PGISFVQKVEIGIYLLDCEAREVYIPRSHQETPVYLGATAGMLRMESEELADRV 148
Db	635	LGNVSLIAKSPFCREBWYIIIGTEEAVFGWALVY-----QTSM-LGALP--KK- 683
QY	149	LDVVERLSINPFD-FOG-ARITIGQEEGAGWTTINYLIGKFSQKTRMFSIVPYETNNO 206
Db	684	ATFGALDGGSSLOVTEENERTHENFNMLRIGSVNHHLSAYSLAGVGLNDAFDRSVH 743
QY	207	ETFPALDLGASITGVTVFPQNGITLESIDNALQFPLVYKDVNVYTHS-F-L-CYKQDA- 261
Db	744	LKRLPNVNSDLIEGLKEMKHPCLNSGVN-GQYISQC-ASSVQGGKRGSGVSIKLVG 801
QY	262	LMQKIAK-DI-QVASNEI-LRDPCEHFHKKVNVSDLYKTPCKRREMPLPQOFEIQC 318
Db	802	APNNGEC--SALAK-NAP-C--AL-P-DG-YPRR-HQFPYAVSGFFVYVFFNINSAF-AS 850
QY	319	IGNVQCHQSLLEFNSTYCPYSCAFANGIFLPLQGDPAFSAFYVMFLNLTSEKVS 378
Db	851	LDDVLEKGRFDCRAMOVARTS-VSP-QPTEIQCFAPIVSLIRGL-YITDK-QI- 904
QY	379	QEKTEBMKRTCAQNPWEIKTSYAGVKEKILSECFGTIISLLDGYFTADSWRH 438
Db	905	IIGS--GSIT-WTLGVAL-L-E--SG-KALSTGLIKSYETLSKNTPIALISILISL 956
QY	439	FIGIKQSDAGWTLGVMILNTMIPABQPISTPLSHSTVYEL-WYLFSLVLFVATIGL 497
Db	957	LL 958
QY	498	IF 499
RESULT	6	
ENTRY	S48859	#type complete
TITLE	nucleoside triphosphatase precursor, chromatin-associated -	
ORGANISM	garden pea	
DATE	13-Jan-1995	#formal_name Pisum sativum #common_name garden pea
ACCESSIONS	29-Sep-1999	#sequence_revision 13-Jan-1995 #text_change
REFERENCE	S65147; S48859	
#authors	S65147	
#journal	Hsieh, H.L.; Tong, C.G.; Thomas, C.; Roux, S.J.	
#title	Plant Mol. Biol. (1996) 30:135-147	
#cross-references	calmodulin-regulated, chromatin-associated NTPase in pea.	
#accession	EMBL:U00000.1	
#molecule_type	mRNA	
#residues	1-455	##label HS2
##cross-references	EMBL:U00000.1; NID:9563611; PIDN:CA63655.1;	
CLASSIFICATION	#superfamily nucleoside triphosphatase chromatin-associated	
KEYWORDS	nucleus	



Matches	22;	Conservative	21;	Mismatches	22;	Indels	2;	Gaps	2;
Dd	228	TRPTGAEGEAFATITLNLHLSRRLEGEDPARCMIDYGVKHCRLNDLAGVVGASQAQIVF	287						
Oy	166	AARITGGEGAYGMVTITNYLLCKRSQSQRTRMFSIYPETENN-QEFL-CALDLGASRQVVF	223						
Dd	288	PLOECTV	294						
Oy	224	VPONOTI	230						
RESULT	11								
ENTRY	JN0620	#type complete							
TITLE	UDP-glucuronosyltransferase (EC 2.4.1.-) 2B-10 precursor - human								
ORGANISM	#formal_name Homo sapiens #common_name man								
DATE	31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 28-May-1999								
ACCESSIONS	JN0620								
REFERENCE	JN0619								
#authors	Jin, C.J.; Miners, J.O.; Lillywhite, K.J.; Mackenzie, P.I.								
#journal	Biochem. Biophys. Res. Commun. 1993;194:496-503								
#title	cDNA cloning and expression of two new members of the human liver UDP-glucuronosyltransferase 2B subfamily.								
#cross-references	MUJID:g9326164								
#accession	JN0620								
#molecule_type	mRNA								
#residues	1-528 #label JIN								
#cross-references	GB:X63359; NID:g516149; PIDN:CMA44961.1; PID:g516150								
CLASSIFICATION	#experimental_source liver								
KEYWORDS	#superfamily glucuronosyltransferase glycoprotein; glycosyltransferase; transmembrane protein								
FEATURE									
1-23									
24-528									
491-508									
66,314,481									
SUMMARY	#length 528 #molecular-weight 60773 #checksum 4903								
Query Match	3.1%; Score 117; DB 2; Length 528;								
Best Local Similarity	24.7%; Pred. No. 1,07e+01;								
Matches	21; Conservative	24; Mismatches	36; Indels	4; Gaps	4;				
Dd	133	KKLKKLDSEFPDVIADAVLPCELLAELENFIV-YSH-SFPSGSFRHGGR-TTP	189						
Oy	303	KRFELTLPFOOFEIGTGIGNYOCHOSILIEFTNSYCPSOCARNGIF-LPPLDGDFGAFS	361						
Dd	190	PSYAPVVMKSLLDOMTFMERVKNNL	214						
Oy	362	AFYFMKFLNTSERVSOEKYTEEM	386						
RESULT	12								
ENTRY	G69998	#type complete							
TITLE	thiodoxin H1 homolog ytpc - Bacillus subtilis								
ORGANISM	#formal_name Bacillus subtilis								
DATE	05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 11-Jun-1999								
ACCESSIONS	G69998								
REFERENCE	A69580								
#authors	Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertello, M.G.; Bessieres, P.; Bolotin, A.; Borcherdt, S.; Boris, R.; Boursier, L.; Brans, A.; Braun, M.; Bridgell, S.C.; Bron, S.; Brouillet, S.; Buschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani, J.J.; Conerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoft, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim,								

```

#submission submitted to the Protein Sequence Database, March 1994
#accession S37995
##molecule_type DNA
##residues 1-919 ##label VA2
##cross-references EMBL:Z28165; NID:g486z88; PID:g486z89; MIPS:YKL165c
##experimental_source strain S288C
REFERENCE S44563
authors Vandewol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hliger F.
#journal Yeast (1994) 10:35-40
#title DNA sequencing of a 3.6.2 kb fragment located between the PAS1 gene and LAP4 loci of chromosome XI of Saccharomyces cerevisiae
#accession S44563
#status nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-919 ##label VA3
##cross-references EMBL:Z26877; NID:g407482; PID:g407483
##experimental_source strain S288C
#note the nucleotide sequence was submitted to the EMBL Data Library, September 1993

GENETICS
#gene SGD:MCD4
##cross-references SGD:S0001648; MIPS:YKL165c
#map_position 111.
SUMMARY #length 919 #molecular_weight 105692 #checksum 6222

Query Match 3.0%; Score 113; DB 2; Length 919;
Best Local Similarity 24.8%; Pred. No. 3.29e+01;
Matches 27; Conservative 35; Mismatches 37; Indels 10; Gaps 9;

Db 422 ENYLEQEAIVITLEPLMOITLEGHLHYLTVMRIRIIVTF-GF-VGMIFPSFIIFLKFSI 479
      I:: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 406 EKYLSEVCSGT-YILSLQGYHF-TADSMEHIFIGKIQGSDAGWT-LGYMLNLTMNI 462
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 480 LENVIDOKASPLSHAVFGSIGILNMILFYOSPPNFMYLLFLPYFW 528
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 463 --PA-EQPLSTPLSHSYVFELMWLFSLVF-TVAILGLLIFFHK-PSYFW 506

RESULT 14
ENTRY
TITLE PXLNPD #type complete
ALTERNATE_NAMES H+-transporting ATPase (EC 3.6.1.35), plasma membrane -
ORGANISM Leishmania donovani
DATE protein-transporing ATPase
formal name Leishmania donovani
30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change
18-Jun-1999
ACCESSIONS A27124
REFERENCE A27124
authors Meade, J.C.; Shaw, J.; Lemaster, S.; Gallagher, G.; Stringer, J.R.
#journal Mol. Cell. Biol. (1987) 7:3937-3946
#title Structure and expression of a tandem gene pair in Leishmania donovani that encodes a protein structurally homologous to eucaryotic cation-transporting ATPases.
#cross-references MUID:88122116
#accession A27124
##molecule_type DNA
##residues 1-974 ##label MEA
##cross-references GB:MI7889; NID:g159291; PIDN:AAA29227.1; PID:g159294
#note the authors translated the codon AGA for residue 352 as Lys
CLASSIFICATION #superfamily Na+/K+-transporting ATPase alpha chain; ATPase
nucleotide-binding domain homology
KEYWORDS ATP; hydrogen ion transport; hydrolase; membrane protein; phosphoprotein
FEATURE
495-670 #domain ATPase nucleotide-binding domain homology #label
ATPN
351 #active-site Asp (aspartylphosphate intermediate)
#status predicted
#length 974 #molecular_weight 107476 #checksum 834
SUMMARY
```



Query Match	3.08;	Score 114;	DB 1;	Length 974;
Post Total ClustLen	30.28;	Need 30.40;	01	

Matches 29; Conservative 24; Mismatches 36; Indels 7; Gaps 5;

Matches 29; Conservative 24; Mismatches 36; Indels 7; Gaps 5;

Db 751 CGSSMLLWIGLEGYSQYENSW--FHRGLAQLPQ-G-KLVTMMYLKISISDFLLFS 806

413 CFSGYILSLLOGY--HFTADSWEHIFIGKIQGSDAGWTLGYMLNLTNMPAEQPLST 470

Db 807 SRTGHHFFFYMP-SPLFCGAIISLVSTMAASF 841

QY 471 PLSHSTYVFLMVESSLVLTVAIGLLIFHKPSYFW 506

## RESULT 15

ENTRY	C70421	#type complete
TITLE	conserved hypothetical protein aq_1392 - Aquifex aeolicus	

ORGANISM	#formal_name	Aquifex	aerolicus
DATE	08-May-1998	#sequence_revision	08-May-1998
		#text_change	

18-Jun-1999  
C70421

REFERENCE	#authors
A70300	
Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.;	

Lenox, A.L.; Graham, D.E.; Overbeek, R.; Sneed, M.A.; Keller, M.; Aujaay, M.; Huber, R.; Feldman, R.A.; Short,

J.M.; Olson, G.J.; Swanson, R.V.  
Nature (1998) 392:353-358

#title	
	The complete genome of the hyperthermophilic bacterium <i>Aquifex aeolicus</i> .

#cross-references MUID:98196666  
#accession C70421

```
#status      preliminary; nucleic acid sequence not shown;
              translation not shown
```

```
##molecule_type DNA
##residues 1-292 ##label AOF
```

```
##cross-references GB:AE000738; NID:g2983801; PID:g2983811; GB:AE000657
##experimental_source strain VF5
```

GENETICS  
#gene aq\_1392

CLASSIFICATION  
#superfamily Methanobacterium thermoautotrophicum conserved  
hypothetical protein MTH382

```
SUMMARY      #length 292  #molecular-weight 33244  #checksum 6083
```

Query Match	2.9%;	Score 11;	DB 2;	Length 292;
Best Local Similarity	32.9%;	Pred. No. 5.70e-01;		

Matches 26; Conservative 22; Mismatches 25; Indels 6; Gaps 6;

D6 163 VLSLLHGVRR-PQFFHSVEIISGLIPVILFSLGRLNFTDMKSDYRTLSAL-FIK- 219  
::|||::| : | : : : : || ||::| : : | ::|

QY 419 ILSLLQGYHFTADSMWH-IHFIGIQSDAGWTLGYMLNTMIPAEQPLSTPLSHSTY 477

Db 220 MELVPLILIV-F-LKIEGL 236  
:||: |:|| |:|:|:

QY 478 VF1MVFSLVLFYVAIGL 496

Search completed: Fri May 5 08:46:49 2000

Job time : 311 secs.



(a) (b) (c) (d) (e) (f)

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: S73813; AAB32152.1; -  
DR EMBL: U87967; AAB47572.1; -  
DR MIM: 601752; -  
DR PROSITE: PS01238; GDA1\_CD39\_NTPASE. 1.  
DR PFAM: PF01150; GDA1\_CD39; 1.  
KW Hydrolyase; Transmembrane; Antigen; Glycoprotein.  
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 17 38 POTENTIAL.  
FT DOMAIN 39 478 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 479 499 POTENTIAL.  
FT DOMAIN 500 510 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 73 73 POTENTIAL.  
FT CARBOHYD 227 227 POTENTIAL.  
FT CARBOHYD 292 292 POTENTIAL.  
FT CARBOHYD 334 334 POTENTIAL.  
FT CARBOHYD 371 371 POTENTIAL.  
FT CARBOHYD 457 457 POTENTIAL.  
SQ SEQUENCE 510 AA; 57964 MW; E403B5C9 CRC32;

Query Match 100.0%; Score 3806; DB 1; Length 510;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MEDKESNVKTEFCSSKNIIILGFTSSIIAVIALAVGLTQNKALPENVKYGIYLDAGSSHT 60  
OY 1 MEDKESNVKTEFCSSKNIIILGFTSSIIAVIALAVGLTQNKALPENVKYGIYLDAGSSHT 60  
DB 61 SLIYKWPAREKNDGVVQVQVEECRVKPGISKFOVKNEIGIYITDCMERAREVTPRSQ 120  
OY 61 SLIYKWPAREKNDGVVQVQVEECRVKPGISKFOVKNEIGIYITDCMERAREVTPRSQ 120  
DB 61 SLIYKWPAREKNDGVVQVQVEECRVKPGISKFOVKNEIGIYITDCMERAREVTPRSQ 120  
OY 61 SLIYKWPAREKNDGVVQVQVEECRVKPGISKFOVKNEIGIYITDCMERAREVTPRSQ 120  
DB 121 HOETPVYLGATGMRLLRMESELDRLVDYVERLSNYPDFOGARITTOEGEGAYGWI 180  
OY 121 HOETPVYLGATGMRLLRMESELDRLVDYVERLSNYPDFOGARITTOEGEGAYGWI 180  
DB 121 HOETPVYLGATGMRLLRMESELDRLVDYVERLSNYPDFOGARITTOEGEGAYGWI 180  
OY 121 HOETPVYLGATGMRLLRMESELDRLVDYVERLSNYPDFOGARITTOEGEGAYGWI 180  
DB 181 TINYLGRFSQKTRWFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240  
OY 181 TINYLGRFSQKTRWFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240  
DB 181 TINYLGRFSQKTRWFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240  
OY 181 TINYLGRFSQKTRWFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240  
DB 241 LYGRKYNYTHSFLCYGKQDALMOKLADIOVASNEILRDPCHPGYKKVNVSDLYKTP 300  
OY 241 LYGRKYNYTHSFLCYGKQDALMOKLADIOVASNEILRDPCHPGYKKVNVSDLYKTP 300  
DB 241 LYGRKYNYTHSFLCYGKQDALMOKLADIOVASNEILRDPCHPGYKKVNVSDLYKTP 300  
OY 241 LYGRKYNYTHSFLCYGKQDALMOKLADIOVASNEILRDPCHPGYKKVNVSDLYKTP 300  
DB 301 CTKRREMTLPQOFELQIGNYQOCHOSTLEFNSTYCPYSQAFNGITPLPLOGDGAF 360  
OY 301 CTKRREMTLPQOFELQIGNYQOCHOSTLEFNSTYCPYSQAFNGITPLPLOGDGAF 360  
DB 361 SAFYVVMFLNTSEKVSQEKYTEMMKKFCAPWEEIKTSYAGVKEKYSEYCFSGTYIL 420  
OY 361 SAFYVVMFLNTSEKVSQEKYTEMMKKFCAPWEEIKTSYAGVKEKYSEYCFSGTYIL 420  
DB 361 SAFYVVMFLNTSEKVSQEKYTEMMKKFCAPWEEIKTSYAGVKEKYSEYCFSGTYIL 420  
OY 361 SAFYVVMFLNTSEKVSQEKYTEMMKKFCAPWEEIKTSYAGVKEKYSEYCFSGTYIL 420  
DB 421 SLLOGYHFTADSWEHIFIGIKIQSDAGWTIGYMLNTNMTPAEOPJSTPLSHSTYVFL 480  
OY 421 SLLOGYHFTADSWEHIFIGIKIQSDAGWTIGYMLNTNMTPAEOPJSTPLSHSTYVFL 480  
DB 481 MYLSVLAFTVAIIIGLIFHKPSYFMKDMY 510  
OY 481 MYLSVLAFTVAIIIGLIFHKPSYFMKDMY 510

RESULT 2  
ID CD39\_MOUSE STANDARD; PRT: 510 AA.  
AC P55772;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE VASCULAR ATP-DIPHOSPHODIROLASE (EC 3.6.1.5) (ATPDASE) (LYMPHOID CELL  
DE ACTIVATION ANTIGEN) (CD39 ANTIGEN).  
GN CD39.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

RN [1]

RP SEQUENCE FROM N.A.

RX MALISZEWSKI C.R., DELESPESE G.J.T., SCHOENBORN M.A., ARMITAGE R.J.,  
RA FANSLAW W.C., NAKAJIMA T., BAKER E., SUTHERLAND G.R., POINDEXTER K.,  
RA BIRKS C., ALPERT A., FRIEND D., GIMPEL S.D., GAYLE R.B. III;  
RT "The CD39 lymphoid cell activation antigen. Molecular cloning and  
RT structural characterization.";  
RL J. Immunol. 153:3574-3583(1994).

CC -1 CATALYTIC ACTIVITY: ATP + 2 H<sub>2</sub>O = AMP + 2 PHOSPHATE.  
CC -1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -1 SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF037366; AAB92259.1; -  
DR MGD: MGI:102805; CD39.  
DR PROSITE: PS01238; GDA1\_CD39\_NTPASE. 1.  
DR PFAM: PF01150; GDA1\_CD39; 1.  
KW Hydrolyase; Transmembrane; Antigen; Glycoprotein.  
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 17 38 POTENTIAL.  
FT DOMAIN 39 478 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 479 499 POTENTIAL.  
FT DOMAIN 500 510 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 73 73 POTENTIAL.  
FT CARBOHYD 226 226 POTENTIAL.  
FT CARBOHYD 291 291 POTENTIAL.  
FT CARBOHYD 333 333 POTENTIAL.  
FT CARBOHYD 428 428 POTENTIAL.  
FT CARBOHYD 457 457 POTENTIAL.  
SQ SEQUENCE 510 AA; 57205 MW; 0570B8FE CRC32;

Query Match 80.3%; Score 3056; DB 1; Length 510;  
Best Local Similarity 76.2%; Pred. No. 0.00e+00;  
Matches 390; Conservative 68; Mismatches 50; Indels 4; Gaps 4;  
DB 1 MEDKESNVKTEFCSSKNIIILGFTSSIIAVIALAVGLTQNKALPENVKYGIYLDAGSSHT 60  
OY 1 MEDKESNVKTEFCSSKNIIILGFTSSIIAVIALAVGLTQNKALPENVKYGIYLDAGSSHT 60  
DB 61 SLIYKWPAREKNDGVVQVQVEECRVKPGISKFOVKNEIGIYITDCMERAREVTPRSQ 120  
OY 61 SLIYKWPAREKNDGVVQVQVEECRVKPGISKFOVKNEIGIYITDCMERAREVTPRSQ 120  
DB 61 SLIYKWPAREKNDGVVQVQVEECRVKPGISKFOVKNEIGIYITDCMERAREVTPRSQ 120  
OY 61 SLIYKWPAREKNDGVVQVQVEECRVKPGISKFOVKNEIGIYITDCMERAREVTPRSQ 120  
DB 121 HOETPVYLGATGMRLLRMESELDRLVDYVERLSNYPDFOGARITTOEGEGAYGWI 180  
OY 121 HOETPVYLGATGMRLLRMESELDRLVDYVERLSNYPDFOGARITTOEGEGAYGWI 180  
DB 121 HOETPVYLGATGMRLLRMESELDRLVDYVERLSNYPDFOGARITTOEGEGAYGWI 180  
OY 121 HOETPVYLGATGMRLLRMESELDRLVDYVERLSNYPDFOGARITTOEGEGAYGWI 180  
DB 181 TINYLGRFSQKTRWFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240  
OY 181 TINYLGRFSQKTRWFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240  
DB 181 TINYLGRFSQKTRWFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240  
OY 181 TINYLGRFSQKTRWFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240  
DB 240 LYGEDYVYTHSFLCYGKQDALMOKLADIOVASNGVLKDCPFGYKGVNVVSELYGTP 299  
OY 241 LYGRKYNYTHSFLCYGKQDALMOKLADIOVASNEILRDPCHPGYKKVNVSDLYKTP 300  
DB 300 CTKRREMTLPQOFELQIGNYQOCHOSTLEFNSTYCPYSQAFNGITPLPLOGDGAF 360  
OY 301 CTKRREMTLPQOFELQIGNYQOCHOSTLEFNSTYCPYSQAFNGITPLPLOGDGAF 360

```

Db 360 SAFEVNDFKFAKNSVISOEKMTEITKNCFSKSWETKTSTPSVKEKYLEYCEFSGAY 419
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 361 SAFEVNDFKFAKNSVISOEKMTEITKNCFSKSWETKTSTPSVKEKYLEYCEFSGAY 418
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 420 ILSLILQGYNEFTDSSWEQIHPFKIKSDNAGWTLGYMLNTNMIPAEQPLSPPLPHSTYI 478
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 419 ILSLILQGYNEFTDSSWEQIHPFKIKSDNAGWTLGYMLNTNMIPAEQPLSPPLPHSTYI 478
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 479 GLMVLFSLLVAVALTGTFTSKSPSYFKKAV 510
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 479 FLMWLFSLVFTVAIIIGLLIFHKPSYFKWDMV 510
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |

RESULT 3
ID YB14.CAEBL STANDARD; PRT; 557 AA.
AC Q21815;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHEICAL 63.1 KD PROTEIN R07E4.4 IN CHROMOSOME X.
GN R07E4.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA MILLER N.;
RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
DR EMBL; U39652; AAA80403.1; -
DR MORPEP; R07E4.4; CE04819.
DR PROSITE; PS01238; GDAL_CD39_NTPASE; 1.
DR PFAM; PF01150; GDAL_CD39; 1.
KM Hypothetical protein; Transmembrane; Hydrolase.
FT TRANSMEM 7 27 POTENTIAL.
FT TRANSMEM 495 515 POTENTIAL.
SQ SEQUENCE 557 AA; 63056 MW; 34CD5D10 CRC32;

Query Match 9.9%; Score 378; DB 1; Length 557;
Best Local Similarity 23.9%; Pred. No. 1.19e-50;
Matches 122; Conservative 134; Mismatches 217; Indels 37; Gaps 30;

Db 12 USAMIFPVI-VFTYVVEAHTSPKVIADDOERSYGVICDAGSTGRLEFVNMISTDSEL 70
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 18 LAIIGFSSIIAVILLAVGLQNKALPEN-VK-YGIYLDAGSSHTSLIYTPAKKENDT 75
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 71 IOIEPVYDNKPVAKKISPGISTGTGTPAQAEYLRPLMELAEKRHPKRYPVVIFFA 130
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 76 GVIVHVE-EGR-V--K-GPGISKFPVKVNEIGIYLTDCMERAREIPRSQHOETPVYIGA 130
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 131 TAGHRLIPDEVLIGOKEAVLKNLKLPTKTSOVULKEHRIITEGKREGIYSMTAVNYA 190
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 131 TAGHRLIPDEVLIGOKEAVLKNLKLPTKTSOVULKEHRIITEGKREGIYSMTAVNYA 190
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 191 LGRNKTATLDPGTPSAHAKRKTQVMIDMGASAGQIAFELPDTSFSSIVENINLCGR 250
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 186 LGRNKTATLDPGTPSAHAKRKTQVMIDMGASAGQIAFELPDTSFSSIVENINLCGR 250
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 251 EDDSLFKFLFTVTFGLGVNEGIRKYEHLMLSLKLDKONGTVIODDCMPLNLRHTVLEN 310
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 241 LYGR--VRYNVTHTSFLCYGKQDALM--QKLAKDIQVANSNEILRDPGFHPGKRVVNSD 295
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 311 G-ENP-VRRGIGNMNTCSNEVKILLN-PESSEYCKA-EAKKCFGAVPAPISILSNEM 366
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |

```

```

QY 296 LYKTPCTRRFEMTLRPFQOFELQIGNVCQHQSLILEFNTSYCPYSGQAFNGIFLPLQOG 355
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 367 -YG-FSEYWTSTHDVILGGO-YDAENIAKKTQOYCSKRWSTIQAESKKQLYPRADERL 423
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 356 DFGAFSAFYFVK--FLNLTSEKVSQEKTEEMKKRCAQPW--E-EIKTS-YAGVKEKYL 409
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 424 RTQCKSAMWTSVLDHGS-VDKTHNKQSVSTIAGQVQALAMITMHMFPPLROSSR 482
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 410 SEYCFSGTYILSLQGHFTADSMHEHFTGKIQSGDAGWTLGYMLNTNMIPAEQPLS 469
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 483 NLIKVEHTSSSESLMAPLEFLSAVCLFVL 512
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 470 TPLSHSTYVFLMWLFSLVFTVAIIIGLLIF 499
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |

RESULT 4
ID YE45_YEAST STANDARD; PRT; 630 AA.
AC P40009;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHEICAL 71.9 KD PROTEIN IN PMI40-PAC2 INTERGENIC REGION.
GN YER005W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA DIETRICH F.S., MULLIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,
RA AVILES E., BERNI A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J.M.,
RA CHONG E., DUNCAN M., GUZMAN E., HARTZEL G., HONICKER-SMITH S.,
RA HYMAN R., KAYSER A., KOMP C., LASHKARI D., LEW H., LIN D.,
RA MASEDALE D., NAKAHARA K., NAMATH A., NORGEN R., OEFNER P., OH C.,
RA PETERL F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOGREN T., SMITH V.,
RA TAYLOR P., WEI Y., YELTON M., BONSTEIN D., DAVIS R.W.;
RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
-----
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
DR EMBL; U18778; AAB64538.1; -
DR PROSITE; PS01238; GDAL_CD39_NTPASE; 1.
DR PFAM; PF01150; GDAL_CD39; 1.
KM Hypothetical protein; Transmembrane; Hydrolase.
FT TRANSMEM 501 517 POTENTIAL.
SQ SEQUENCE 630 AA; 71851 MW; 68CD0D15 CRC32;

Query Match 9.5%; Score 360; DB 1; Length 630;
Best Local Similarity 28.4%; Pred. No. 5.89e-47;
Matches 106; Conservative 93; Mismatches 137; Indels 37; Gaps 28;

Db 52 IHOEKDWTFKLNPGLSPKPEKRODAYKSHIKPLDFAKNIPESHWSCSPFYQIATAGMR 111
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 78 VHQVECKRYK-GPQISKFPVKVNEI-GIYLTDCMERAREIPRSQHOETPVYIGA 135
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 112 LLPDIOSSIIIDGLCKGKHP-AEFLVEDCSAQIOIVIDGETGLGYGLNLYLGHFND- 169
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 136 LIRNESEE-LADRYLDVYERSLNVPF-DFGA-RITIGQEGAGVGTIYLLGKTSQK 192
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 170 ---YN--P-EYSDHFTFGMDMGASQIOIAPAPDSGEIARHARDIATIFLRVNGDLQK 223
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 193 TRWPSIVYETNNQETFGALDYGASTQVFPVPOQOT-IESP-DN-ALQF-R-LYGR-D- 245
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 224 MDVFSVWIGLFGANQARRRYLAOLINFLPENT-ND-YENDQFSRR-NLNDQCMRGSGSTD 280
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |

```

QY	246	YNYVTHSLFCGKQDALMQKLAKDIOQVANSNBLDPCPHNYKKVYVWVSDLYKPR-CTKR	304
		281 FEFKDT--PHIAGSGNYEGCTKSTIYPLLLKNMPCQDDEPCLFNGVNAHPRIDFANDKFTGT	338
		305 FEMTLPEQFOEFIEQIGNGQQCHOSLIELEFNSTY-CPYSGCAFNGIIFPPIDY-G-D-PGAF	360
Db	339	SEYWTANDYKELGE-YNFDKFSKSLREFPNSWVQTQIANSDDGVNSIPENTLAKACF	397
QY	361	SAFFYVMK-FENLTSEKVSQEKVTEMMKKFCQAPWEI-----RTSYAGVGEKYLSEYCF	414
Db	398	GNWVNLNITHEGF	410
QY	415	SGYIILSLLOGY	427

RESULT	5	STANDARD:	PRT:	485 AA.
ID	YVAE.CAEEL			
AC	018411:			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	HYPOTHETICAL 54.3 KD PROTEIN C33H5.14 IN CHROMOSOME IV.			
GN	C33H5.14.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Secernentea; Rhaditida; Rhaditidae;			
OC	Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2.			
RA	BRADSHAW H., STELLYES L.;			
RL	Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.			
CC	-1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; U41007; AAA82272.1; -			
DR	WORMPEP; C33H5.14; CE04157.			
DR	PROSITE; PS01238; GDAL_CD39_NTPASE; 1.			
DR	PFAM; PF01150; GDAL_CD39_1			
FT	Hypothetical protein; Transmembrane; Hydrolase.			
FT	TRANSMEM 439 439 POTENTIAL			
FO	SEQUENCE 485 AA; 54309 MW; 00659F2B CRC32;			

Query Match	8.58;	Score 324;	DB 1;	Length 485;
Best Local Similarity	24.98;	Pred. No. 1.18e-39;		
Matches	116;	Conservative 113;	Mismatches 196;	Indels 40;
			Gaps 34	

Db 21 NIKKGVACDSSGSTRFEVYTLPLSGGLTIDITLHESPPVKKYTPGLSSGDPPEO 80  
 Qy 45 ENVKKGYLVDGSSHTSLYIKW-P-AEKENDTG-VYHQVECKVW-GPGLSKVQKAVNE 100Y  
 Db 81 VVEYLPPLFAEENHPIYEQLGETDLLIFATGRLLPEAKODAIKKNLNGKSVATLR 140  
 Qy 101 IGYLTQCMERAKARVYIPRSQHOETPVYLGATGAKRLLRMESEELADVLVDVERSLNYP 160Y  
 Db 141 VSDSNIRIIOAMEGISWIAVNTIIRREPDKENS-KYGMIDMGASVOJAFELANE-KE 198Y  
 Qy 161 FDFOGARLTQEGEAGWITINTLLGFSOKTWFESTIVPEYETNNQETFGALDAGASTQ 220Y  
 Db 199 -SYNGGN-VYEINGISJETNEDY-K-YKISTPPLGAGANGL-KKYENSL-VKSGNS-N 251Y  
 Qy 221 VTFVPOQOTIESPNDALQFRL-YEKDYNVYTHSFLCYGKDQALMQKLAKDIQVANSIELR 279Y  
 Db 252 DSCSPRGKINRLI-GE-FTVANGTEMBDYCA-QVYSSL-TGD-KA-QPSC-P-NPL-CF 299Y  
 Qy 280 DPCHPGKGVKAVNSDYKTPCTIRREFETLTFPOOFLQIGANQOQCHSLLELFNYSYCP 339Y

Db 300 LRNVIAIPSVNLSTVOL -G-ESEKYTTSSNGSGGEHYQ-KFDEYKTKCQKMDNDIOD 356

Qy 340 YSQAENMELFPLPDGDGASAFAYFNKFLINTJSEKVSQEKVEMKKKCAQPMEEI-- 397

Db 357 GFKRREFPNADIERTGNTCFKAAMVTSVLHGFN-VDTKHLFEQSVYKIGEEKONALGA 415

Qy 368 --KTS-YAGVEKYLSEKCFSGTITLSTLLDGGHFTADSWENHIFICKIGSDGKWTIGY 454

Db 416 MLVHSKDKLFNLEQLEVAOSTQOISNF-SFF-VILITVLAVAL 458

Qy 455 ML-NITNM-IPAEQPLSTPLS-HSTYVFAMLESLVETVAITL 496

RESULT ID	6	STANDARD:	PRT:	454 AA.
AP	APY_SQUTU			
AC	P80595: 043164:			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	APYRASE PRECURSOR (EC 3.6.1.5) (ATP-DIPHOSPHATASE) (ADENOSINE DIPHOSPHATASE) (ADPASE) (ATP-DIPHOSPHOHYDROLASE).			
GN	ROPO1.			
OS	Solanum tuberosum (Potato).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.			
RP	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 59-160; 236-253 AND 332-345.			
RC	TISSUE-TUBER:			
RX	MEDLINE; 96158985.			
RA	HANDA M., GUDIDOTTI G.;			
RT	"Purification and cloning of a soluble ATP-diphosphohydrolase (apyrase) from potato tubers (Solanum tuberosum).";			
RT	Biochem. Biophys. Res. Commun. 218:916-923(1996).			
RL	[2]			
RP	SEQUENCE OF 42-54; 68-95 AND 236-253.			
RC	STRAIN-CV. DESIREE;			
RX	MEDLINE; 96355615.			
RA	VASCONCELOS E.G., FERREIRA S.T., DE CARVALHO T.M.U., DE SOUZA W., KETTLUNG A.M., MANGILA M., VALENQUEIRA M.A., VERJOVSKI-ALMEIDA S.;			
RT	"Partial purification and immunohistochemical localization of ATP diphosphohydrolase from Schistosoma mansoni. Immunological cross-reactivities with potato apyrase and Toxoplasma gondii nucleoside triphosphate hydrolase.";			
RT	J. Biol. Chem. 271:22193-22145(1996).			
RL	[1]			
CC	-1- FUNCTION: CATALYZES THE HYDROLYSIS OF PHOSPHOANHYDRIDE BONDS OF NUCLEOSIDE TRI- AND DI-PHOSPHATES.			
CC	-1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.			
CC	-1- COFACTOR: CALCIUM.			
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED (PROBABLE).			
CC	-1- PFM: THE N-TERMINUS IS BLOCKED.			
CC	-1- SIMILARITY: BELONGS TO THE GAI1 / CD39 NTPASE FAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb.ch/announce/">http://www.isb.ch/announce/</a> or send an email to <a href="mailto:license@isb.ch">license@isb.ch</a> ).			
CC	-----			
DR	EMBL: U58597; AAB03720.1; -			
DR	PROSITE; PS01238; GDA1_CD39_NTPASE; 1.			
DR	PFAM; PF01150; GDA1_CD39; 1.			
KW	Hydrolase; Transmembrane; Calcium; Signal.			
FT	SIGNAL	1	30	POTENTIAL.
FT	CHAIN	31	454	APYRASE.
FT	TRANSMEM	426	446	POTENTIAL.
FT	CARBOHYD	151	151	POTENTIAL.
FT	CARBOHYD	262	262	POTENTIAL.
FT	SEQUENCE	454 AA;	50041 MW;	D6FAEA89 CRC32;







```

RT Toxoplasma gondii."
RL J. Biol. Chem. 270:11391-11397(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-RH;
RX MEDLINE: 95050750.
RA BERMEJES D., PECK R.R., AFIFI M.A., BECKERS C.J.M., JOINER K.A.;
RT "Randomly repeated genes encode nucleoside triphosphate hydrolase
RT isoforms secreted into the parasitophorous vacuole of Toxoplasma
RT gondii."
RL J. Biol. Chem. 269:29252-29260(1994).
CC -1- FUNCTION: MAY PERFORM AN IMPORTANT PROCESSING STEP IN THE
CC CONVERSION OF HIGH ENERGY NUCLEOTIDES PRIOR TO UPTAKE BY THE
CC PARASITE AND MAY CONTRIBUTE TO INTRACELLULAR SURVIVAL AND
CC VIRULENCE. NTPASE-1 HAS A SPECIFIC ACTIVITY 4.5-FOLD HIGHER THAN
CC NTPASE-II IN HYDROLYSIS OF ATP. THE PRIMARY DIFFERENCE BETWEEN
CC THESE ISOZYMES LIES IN THEIR ABILITY TO HYDROLYZE NUCLEOSIDE
CC TRIPHOSPHATE VERSUS DIPHOSPHATE SUBSTRATES. WHILE NTPASE-II
CC HYDROLYZES ATP TO ADP AND ADP TO AMP AT ALMOST THE SAME RATE,
CC NTPASE-I HYDROLYZES ADP TO AMP AT A MUCH SLOWER RATE (0.7% OF THE
CC RATE FOR ATP).
CC -1- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + PHOSPHATE.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: SECRETED; FOUND IN HOST CELL PARASITOPHOUS
CC VACUOLE.
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L39078; AAA89203.1; -
DR EMBL: U96965; AAC80188.1; -
DR PROSITE: PS01238; GDAL_CD39_NTPASE; 1.
DR PFAM: PF01150; GDAL_CD39; 1.
KW Hydrolase; Multigene family; signal.
FT SIGNAL 1 25
FT CHAIN 26 628 NUCLEOSIDE-TRIPHOSPHATASE I.
FT CARBOHYD 432 432 POTENTIAL.
FT SEQUENCE 628 AA; 69159 MW; DA2A1577 CRC32;
SQ
Query Match 3.3%; Score 125; DB 1; Length 628;
Best Local Similarity 32.8%; Pred. No. 1,31e-03;
Matches 22; Conservative 20; Mismatches 23; Indels 2; Gaps 2;
Db 228 TRPITGAEGLFAFTITLNLHSRRIGEDPARCMIDEGYKQKQNDLAGVYEGGASQAQIVF 287
:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 166 ARIITGEGAGAVGITIVYLKGFQSKTRMFESIYVETNN-QEFP-GALDIGASTGYTF 223
Db 288 PLOEGTV 294
:|:|:|
QY 224 VPQNGTI 230

```

```

RA MEDLINE: 93326164.
RX JIN C.-J., MINERS J.O., LILLYWHITE R.J., MCKENZIE P.I.:
RT "CDNA cloning and expression of two new members of the human liver
RT UDP-glucuronosyltransferase 2B subfamily."
RL Biochem. Biophys. Res. Commun. 194:496-503(1993).
CC -1- FUNCTION: UPPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
CC ENDOGENOUS COMPOUNDS.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
CC BETA-D-GLUCURONOSIDE.
CC -1- SUBCELLULAR LOCATION: MICROSOMAL.
CC -1- SIMILARITY: BELONGS TO THE UDP-GLUCOSYLTRANSFERASE FAMILY.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC DR EMBL: X63359; CAA44961.1; -.
CC DR PIR: JN0620; JN0620.
CC DR MIM: 600070; -.
CC DR PROSITE: PS00375; UDPGT: 1.
CC DR PFAM: PF00201; UDPGT: 1.
CC KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane; Signal;
CC Multi-gene family; Microsome.
CC FT SIGNAL 1 23 BY SIMILARITY.
CC FT CHAIN 42 528 UDP-GLUCURONOSYLTRANSFERASE 2B10.
CC FT TRANSMEM 492 512 POTENTIAL.
CC FT CARBOHYD 66 66 POTENTIAL.
CC FT CARBOHYD 314 314 POTENTIAL.
CC FT CARBOHYD 481 481 POTENTIAL.
CC SQ SEQUENCE 528 AA; 60774 MW; 63F280A6 CRC32;
CC -----
Query March 3.1%; Score 117; DB 1; Length 528;
Matches Similarity 24.7%; Pred. No.1,77e-02;
Matches 21; Conservative 24; Mismatches 36; Indels 4; Gaps 4;
-----
Db 133 KIKMKIKQESRPDIYFAADAYLPCGELLAELFNIPYV-YSH-SFSPGYSPERHSGGF-IFP 189
OY 303 KRFEFTLPQGEIIGIGYQOCHOSITLFTFTSYCPYQCAFNGIF-LPPIQSGFGAHS 361
Db 190 PSYVEPVMSKSIDQMTFEMERVKNML 214
OY 362 AFYEVAKFLNLTSEKVSQEKYTEMM 386
-----
RESULT 13
ID ACHX_ONCYO STANDARD; PRT; 436 AA.
AC P54247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1986 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE ACETYLCHOLINE RECEPTOR PROTEIN, NON-ALPHA CHAIN (FRAGMENT).
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Spirurida;
OC Filarioidae; Onchocercidae; Onchocerca.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94299155.
RA AJUH P.M., EGWANG T.G.;
RT "Cloning of a cDNA encoding a putative nicotinic acetylcholine
RT receptor subunit of the human filarial parasite Onchocerca
RT volvulus."
RL Gene 144:127-129(1994).
CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; L20465; AAA21823.1; -.
CC DR EMBL; L12543; AAA29415.1; -.
CC DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
CC DR PfAM; PF00065; neur_chan; 1.
CC KW Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane;
CC Multigene family.
CC FT NON_TER 1
CC FT DOMAIN 1 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 196 219 POTENTIAL.
CC FT TRANSMEM 227 245 POTENTIAL.
CC FT TRANSMEM 261 280 POTENTIAL.
CC FT DOMAIN 281 404 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 405 423 POTENTIAL.
CC FT DISULFID 89 103 BY SIMILARITY.
CC FT CARBOHYD 62 62 POTENTIAL.
CC FT CARBOHYD 140 140 POTENTIAL.
CC SQ SEQUENCE 436 AA; 51340 MW; 136781FD CRC32;
CC -----
CC Query Match 3.0%; Score 116; DB 1; Length 436;
CC Best Local Similarity 38.5%; Pred. No. 2,44e-02;
CC Matches 20; Conservative 17; Mismatches 9; Indels 6; Gaps 6;
CC -----
Db 241 YLL-LKDIIPAT-STALPL-FCKYLLFTFMIMVSVL-VTVSLNLRHRRPS 288
Oy 454 YMLNTMTPAEOPLSTPLSHSTYV-FLMWLFSLVFLVAIIGL-LIFHKPS 503
CC -----
CC RESULT 14
CC ID YK05_YEAST STANDARD; PRT; 919 AA.
CC AC P36051.
CC DT 01-JUN-1994 (Rel. 29, Created)
CC DT 01-JUN-1994 (Rel. 29, Last sequence update)
CC DT 01-OCT-1994 (Rel. 30, Last annotation update)
CC DE HYPOTHEETAL105.7 KD PROTEIN IN TPk3-PIR1 INTERGENIC REGION.
CC GN YKL165C OR YKL619.
CC OS Saccharomyces cerevisiae (Baker's yeast).
CC OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
CC OC Saccharomycetaceae; Saccharomycetes.
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-S288C;
CC RX MEDLINE; 94378720.
CC RA VANDENBOL M., BOLE P.-A., DION C., PORTETELLE D., HILGER F.;
CC RT "DNA sequencing of a 36.2 kb fragment located between the FAS1 and
CC YAP loci of chromosome XI of Saccharomyces cerevisiae.";
CC YL Yeast 10:535-540(1994).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; Z26877; CA81489.1; -.
CC DR EMBL; Z28165; CA82007.1; -.
CC DR PIR; S37786; S37786.
CC DR PIR; S44563; S44563.
CC KW Hypothetical protein.
CC SQ SEQUENCE 919 AA; 105693 MW; 73C93BED CRC32;
CC -----
CC Query Match 3.0%; Score 113; DB 1; Length 919;
CC Best Local Similarity 24.8%; Pred. No. 6,24e-02;

```









[illegible]

D	b		65	NGTGIYQVVECTYANGSGISSYADDPDAGASLPCPLDLMAMAVIPEEQMOTPTTLGATA	124
O	y		73	:::    ::    :    :	132
D	b		125	GMRLLREONSTKAOVAEFSKAIREPVDGRGAOLITGNEBGSFGMITVNLLETFL-K	183
O	y		133	GMRLLRMESELDARVLADVRSISNTPDFQCARLITGOEBATGMITINILLOKESOK	192
D	b		184	FS-FA-GKMEHPONTTEVLGLALDLGASTOYTEOP-GVTLEDNTSIFRLXTNSTLYTH	240
O	y		193	TNRMSIYPYE-TNNQEIRFALLDLGASGTQTFVPQNQTIESPDMLQPRLVSKDVNYTH	251
D	b		241	SYLCYGQIOIA-SKRLLMALHODGSYVONISHPCYKGYRRITTAETIDSPCVPPSMS	299
O	y		252	SFLCYQRKDQALMOKLARDI-QVASN-EILLDPCEPHGYKKVVNSDLYXTPCKTRFEWTL	309
D	b		300	PAOLLTYTGNGNPACPTALIKLEPNLTCCANRCRGSDGYOYRPVNGQAFAPGFYTTSF	359
O	y		310	PFOQFEILOGIGNTQQCHOSILELFNFSTYCYSOCANGLFPLRLGDGCAFSAFYVKPF	369
D	b		360	LNLGGQ-SLSHVATYWDCCNKNWSLEVTEFPONKE-HLHTYCVVGXYLIILLVDGYKF	417
O	y		370	LNLISEKYSQSKAYTEMKMKKRCACAPWEIEKTSYAGVXEKLSYCSGYILLSLLOGHF	429
D	b		418	DEHTWSNIHSQKAGNADICMTLGFMILNTNMIPTE-ALEHWKGHEPSIMAGAT-SFTVL	475
O	y		430	TADSMEHIEHFQIGQSDDACMTLGMYMLNTNMPABQPLSTPLSHSYVFLLVLPFLSVLF	489
D	b		476	AI-VAGIVAILLOC-FWKSK	493
O	y		490	TVALITGLITFHFKPSYFWKDM	509
R	E	S	RESULT	5	
I	D		ID	055026	PRELIMINARY; PRT: 495 AA.
A	C		AC	055026;	035928;
D	T		DT	01-JUN-1998	(TREMBrel, 06, Created)
D	T		DT	01-JUN-1998	(TREMBrel, 06, last sequence update)
D	T		DT	01-NOV-1999	(TREMBrel, 12, last annotation update)
D	E		DE	ECOTO-ATPASE	(EC 3.6.1.5).
G	N		GN	CD39L1.	
O	S		OS	Mus musculus (mouse).	
O	C		OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
O	C		OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
R	N		RN	[1]	
R	C		RC	SEQUENCE FROM N.A.	
R	C		RC	TISSUE-LIVER.	
R	X		RX	MEDLINE; 98288263.	
R	A		RA	GAO L., DONG L., WHITLOCK J.P. JR.;	
R	T		RT	"A novel response to dioxin. Induction of ecto-ATPase gene	
R	T		RT	expression.";	
R	L		RL	J. Biol. Chem. 273:15358-15365(1998).	
R	N		RN	[2]	
R	P		RP	SEQUENCE FROM N.A. (SHORT FORM).	
R	C		RC	TISSUE=EMBRYO;	
R	X		RX	MEDLINE; 97419269.	
R	A		RA	CHADWICK B.P., FRISCHAUF A.M.;	
R	T		RT	"Cloning and mapping of a human and mouse gene with homology to ecto-	
R	T		RT	ATPase genes.";	
R	L		RL	Mamm. Genome 8:668-672(1997).	
R	C		RC	-1- FUNCTION: HYDROLASES EXTRACELLULAR ATP TO AMP.	
R	C		RC	-1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.	
R	C		RC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POSENTIAL).	
R	C		CC	-1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; A LONG FORM (SHOWN HERE) AND A	
R	C		CC	SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.	
R	C		CC	-1- INDUCTION: BY DIOXIN.	
R	C		CC	-1- PM: HAS PROBABLY VARIOUS DISULFIDE-BONDS (BY SIMILARITY).	
R	C		CC	-1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPEASE FAMILY.	
R	R		DR	EMBL; AF042811: AAC24347..1.. -.-	
R	R		DR	EMBL; U91511: AAB81014.1.. -.-	
R	R		DR	MGD; MG1:1096865; Cd39l1.	
R	R		DR	PROSITE; PS01238; GDAL_CD39_NTPEASE; 1.	
R	R		DR	PFAM; PF01150; GDAL_CD39; 1.	

Accession	Protein	Score	DB	Length
KM	Hydrolase; Transmembrane; Antigen; Glycoprotein; Alternative splicing.	37.58;	DB 11;	Length 495;
FT	DOMAIN 1	43.78;	Pred. No. 1,42e-258;	
FT	TRANSEM 2	103;	Mismatches 152;	Indels 21; Gaps 19
FT	DOMAIN 25			
FT	DOMAIN 26			
FT	TRANSEM 463			
FT	DOMAIN 467			
FT	DOMAIN 480			
FT	VARSPPLIC 130			
FT	VARSPPLIC 133			
FT	CARBOHYD 64			
FT	CARBOHYD 129			
FT	CARBOHYD 294			
FT	CARBOHYD 319			
FT	CARBOHYD 378			
FT	CARBOHYD 443			
SO	SEQUENCE 495 AA; 54310 MM; B9CE5702 CRC32;			

D	b		7	SUPLPALLAAGLAGLLILCPV-PODUREPALYGGVILADGSSHTSMFYKKPADEND	66
Oy			16	NILAIIIGESSI-INAVIALLANGLTQNKALPENNVKTYGIYLDAGSSHTSLYIKKPAEKEND	74
D	b		66	TGIVGOHSSCDVRGGGSISSYANDPSRAGOSLVECLFEQLALHDY-PKDRYSTAPRYLGATAG	1244
Oy			75	TGVVHOUEECRVKRGPKSGFKFQVKNEIGIYLTDCCMERA-REVIPRSQHOFEPYLAGTAG	1338
D	b		125	MRLNLTSPEPTAYVLEAVYOTLTRYPEDRGARILSGODEGVYGVYANTYLENF-KY	1833
Oy			134	MRLRMESSELADVYLVERLSLNYPEPDGARIITNGOEBGAAGWMTYNYLLGKFESQKT	1933
D	b		184	GM-VGRMIRPRKGTGLGAMDGCASOTIEFTETSPS-EDDPNENHYLYGONYRVYTHSF	2400
Oy			194	RWEISIVEPETNNQTFEGALDLGGASTOYTTPYRONQTIESPDNALQFELYLKDYKNVTIHSHF	2533
D	b		241	LGYRGDOVL-ORLLASALQI--HNF-H-PCWPKGYSTOVLLREYOSRPCMGORPOTFNS	2955
Oy			254	LCYGKDQALMQKL-AKDIOVASNELLRDPCRHPPKQKVUNVSUDLYKRPCK-RPEMTL-P	3100
D	b		296	SATYSLSGSTMAALCRDLYVSLGFNISSCFPSQCSFNQVPDPVAAGNTIASATYYVDLF	3555
Oy			311	PQOFEIOGIGNYOOCHOSILEFTMSICPYSCAFNGLFLPRLDGGFGAESAIFYPWKFL	3700
D	b		356	KTVGGLPRTGKOLEDATETTCNCTMAELDARVVGQOTR-LPDVCYAMFIHOLLNRGYS	4140
Oy			371	N-LISEKVSQEK-VTEMMKKFCAPQWEEIKTSTYAGVAKELISCEFSGTYIISLLDGTH	4288
D	b		415	FDESRFGVVEEKKRADTAVGMTLGYMLNTNLIPADLPGLRKGRTHRSSVVALLLFT-V	4733
Oy			429	FTADSWEHIHFICKIQSDAGWTILGYMLNTNMIPAQP-TSPLSHSYTVFLMVLEFSLV	4870
D	b		474	LILAALVLL 483	
Oy			488	LFVTAIIGLL 497	
RESULT			6		
ID			035795	PRELIMINARY;	PRT; 495 AA.
AC			035795;		
DT			01-JAN-1998	(TREMBLrel. 05, Created)	
DT			01-JAN-1998	(TREMBLrel. 05, last sequence update)	
DT			01-NOV-1999	(TREMBLrel. 12, last annotation update)	
DE			ECTO-ATPASE	(EC 3.6.1.5).	
OC			Rattus norvegicus	(Rat).	
OC			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
OC			Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
RN			[1]		
RC			SEQUENCE FROM N.A.		
RC			STRAIN=SPRAGUE DAWLEY; TISSUE=BRAIN;		
MD			MEDLINE; 98031057.		

RA	KGEBL B. BRAUN N., HEINE P., MALISZEWSKI C.R., ZIMMERMAN H.;
RT	an ecto-ATPase and an ecto-ATP diphosphohydrolase are expressed in
RT	ret brain.";
RL	Neuropharmacology 36:1189-1200(1997).
CC	-1- FUNCTION: HYDROLYSES EXTRACELLULAR ATP TO AMP.
CC	-1- CATALYTIC ACTIVITY: ATP + 2 H(2O) = AMP + 2 PHOSPHATE.
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC	-1- TISSUE SPECIFICITY: HIGHEST LEVELS IN HEART, KIDNEY, AND SPLEEN.
CC	STRONG SIGNALS IN THYMUS, LUNG, SKELETAL MUSCLE, AND BRAIN. VERY
CC	WEAK SIGNAL IN LIVER.
CC	-1- PTM: HAS PROBABLY VARIOUS DISULFIDE-BONDS.
CC	-1- SIMILARITY: BELONGS TO THE GDAI / CD39 NTPASE FAMILY.
DR	EMBL: Y11835; CAA12533.1; -
DR	PROSITE: PS01238; GDAI_CD39_NTPASE; 1.
DR	PFAM: PF01150; GDAI_CD39; 1.
KW	Hydrolase; Transmembrane; Antigen; Glycoprotein.
FT	DOMAIN 1 4 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 5 25 POTENTIAL.
FT	DOMAIN 22 25 POLY-LEU.
FT	DOMAIN 26 462 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 463 483 POTENTIAL.
FT	DOMAIN 467 470 POLY-LEU.
FT	DOMAIN 484 495 CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD 64 64 POTENTIAL.
FT	CARBOHYD 129 129 POTENTIAL.
FT	CARBOHYD 294 294 POTENTIAL.
FT	CARBOHYD 306 306 POTENTIAL.
FT	CARBOHYD 319 319 POTENTIAL.
FT	CARBOHYD 378 378 POTENTIAL.
FT	CARBOHYD 443 443 POTENTIAL.
SEQ	SEQUENCE 495 AA; 54389 MW; 24F8B16C CRC32;
Query Match	37.4%; Score 1422; DB 11; Length 495;
Best Local	Similarity 44.0%;
Matches 213; Conservative	98; Mismatches 154; Indels 19; Gaps 17

D0	13	L	L	A	A	G	I	T	G	L	-	L	C	P	-	I	Q	V	R	E	P	P	A	L	K	T	G	I	V	D	A	G	S	H	T	S	M	E	V	T	K	M	P	A	K	N	D	T	G	T	G	70					
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	71					
QY	20	I	L	F	S	S	I	A	I	A	T	A	L	A	G	L	T	O	N	K	A	L	E	P	N	K	T	G	I	V	D	A	G	S	H	T	S	L	T	I	Y	K	M	P	A	K	N	D	T	G	79						
D0	71	O	H	S	S	C	D	O	G	G	I	S	S	A	N	D	E	S	K	A	G	O	S	T	I	V	C	E	D	O	A	L	R	D	-	P	R	R	N	A	S	P	L	Y	G	A	T	A	G	R	P	N	12				
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	13					
QY	80	Q	V	E	C	R	K	F	G	I	S	K	V	E	K	O	K	N	E	I	G	I	T	L	C	M	E	R	A	-	R	E	V	I	P	S	O	H	E	T	P	V	I	G	A	T	A	G	R	L	I	R	13				
D0	130	L	I	S	P	E	A	T	A	R	V	E	A	N	T	O	T	L	Q	V	E	P	F	E	D	R	G	A	R	I	S	G	O	D	E	G	V	E	N	T	A	N	Y	L	L	E	N	F	I	-	K	Y	G	-	-	V	18
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	19					
QY	139	M	E	S	E	L	A	D	R	L	D	V	E	R	S	L	S	N	P	E	F	D	G	A	R	I	I	T	G	O	D	E	G	A	Y	G	M	I	T	I	N	Y	L	L	K	F	S	O	K	T	H	W	F	S	I	19	
D0	187	G	R	N	I	P	R	K	G	T	L	G	A	M	I	D	G	A	S	T	O	I	T	E	T	T	S	P	-	E	D	P	N	E	V	A	L	R	L	G	H	I	R	Y	T	H	S	F	L	C	T	Y	R	24			
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	25					
QY	199	V	P	E	T	N	O	E	F	G	A	L	D	I	G	G	A	S	T	O	V	T	E	V	P	P	O	N	O	T	E	S	D	N	A	L	O	F	R	L	G	K	O	N	Y	T	H	S	F	L	C	T	Y	R	25		
D0	246	D	O	L	L	R	L	S	A	L	O	-	H	R	-	H	-	P	C	M	P	K	T	S	T	O	V	L	O	E	V	O	S	P	C	M	O	G	R	P	A	N	G	S	A	I	V	L	30								
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	31					
QY	259	D	O	A	L	M	O	K	I	A	D	I	O	V	A	S	N	E	L	R	D	P	C	H	N	P	K	K	V	A	N	S	D	L	Y	K	P	C	K	-	R	E	M	T	L	P	P	O	F	-	B	I	31				
D0	302	S	G	S	N	A	T	R	C	D	L	S	R	E	F	I	S	S	C	P	F	O	C	S	F																																



[illegible]

QY	256	YKDDALMOKLAKDIOVAS - NEILRDPCHFGYKKVYVAVSDLYKTPCKRREMTLPQO -	313
Db	299	VVSGTGNGLCAVHVNKLDFPTSCSFHSCFDFGQDEVSNGFIASAFYTYDFIRTV	358
QY	314	FEIIGIGVVOOCHOSILEFNTSYCPYSQCAFNGFIETLPLOGDGAFAFYFNMKFLUL	373
Db	359	MRRPHSBDLKADETICATSMNMLYOK - APRLEKRLPDVCARSTPYLLITGYNPN	417
QY	374	SKR - V - SOEKVEMMKKFCADQAPWEIKTSTYGVKREKYLSEYCFSGTYLLSLLOGYHETA	431
Db	418	RFPPIAOKKRGFESIGMALGYMLNTNMIPAOEPAHSRMLYNNYVLLILLEVITLT	477
QY	432	DSWEIHIFIGKIQSGDAGTLCGYMLNTMIPAEPLS - TPLSHSTYFVFWLVSLEYLT	490
Db	478	ALLTAVYLLRS 489	
QY	491	VAITGLIFHKP 502	
RESULT	8		
ID	Q9Y5L3	PRELIMINARY;	PRT; 495 AA.
AC	Q9Y5L3;		
DT	01-NOV-1999 (TREMBLrel. 12, Created)		
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)		
DT	01-NOV-1999 (TREMBLrel. 12, Last annotation update)		
DE	ECOT-ATPASE (EC 3.6.1.3).		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
OC	Eutheria; Primates; Catarrhini; Homnidae; Homo.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MATEO J., HARDEN T.K., BOYER J.L.;		
RT	"Functional expression of a cDNA encoding a human vascular endothelial		
RT	ecto-ATPase.";		
RL	Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL, AF144748; AAD40239.1; --		
KW	Hydrolase.		
SO	SEQUENCE 495 AA; 53665 MW; 60C0EB61 CRC32;		
Query Match	35.5%;	Score 1353;	DB 4; Length 495;
	Best Local Similarity 40.3%;	Pred. No. 8,93e-243;	
Matches 197;	Conservative 117;	Mismatches 158;	Indels 17; Gaps 16
Db	6	RLSLPPLLAAGLAGLLLCVPRDVR - PPAIKYGIIVLDAGSSHTSMFTYKPPADKEN	64
QY	15	KILAILGPRSS - IAVIALLVAGLQNKALDPENVKYGIIVLDAGSSHTSLTYKPPAEKEN	73
Db	65	DTGIGQHSSCDYPGCGGISSTADNPGASQSLVGCLEALDQDVERKHAITPLYLGATAG	124
QY	74	DGVVHQAWECEKVPKPGISKFEVQKNEIGIYLLDCEMERARAVIPRSQHOETPVYLGATAG	133
Db	125	MLLMLTNPBEASTVLMVTHITLDYPPDPFGARILISQEEGVGVWYVIANLENF - KY	183
QY	134	NMLLMESEELADRLVDIVERSLSWTPDPFGARILITQEEGAYGWTINILLLKFSOKT	193
Db	184	GNVGR - WFRPRK - GTLGAMDLCGASTQITFETTSBA - BDRASEVOLHLGQHYRYTHSF	240
QY	194	RMFSTVPEPTNNGQETFGALDLGGASTQYTFVPQNGTIESPDNALQFRLYKGDVNVYTHSF	253
Db	241	LCYGDQVYL - ORLLASA - LQTHGF - H - PCWPRGSTOVLCDVYQSQCTMAORPONENSS	296
QY	254	LCYGDQALMOKLADIOVASNEILRDCPFHGYKKVYVAVSDLYKTPCK - --REFEMTLPF	311
Db	297	ANVSLSGSSDPLLCGLDYSGLSFSSSCPFSCFSGNGVGPVAGVAFSAFETYVFLR	356
QY	312	QOPEFIOGIGNTVOOCHOSILEFNTSYCPYSQCAFNGFIETLPLOGDGAFAFYFNMKFLN	371
Db	357	TSMGPIVATLLOLEAAAVNVCNQTMAQOLAVRPGORAR - LADYACAGAFYVOLLSSRGGF	415
QY	372	LTFSE - KVSQ - EKVTMEMKKFCQAPHEBELIKTSGYAKKEKYLEYSEFSGYIILSLLOGYH	429
Db	416	DERAFGVIYFOKKAADTVAGMALGYMLNTLIPADPPGLKRGTDSESSVVLILFLPASAL	475

OY 430 TADSWEHIFIGKIOGSDAGWTGLGMLNTNMIPAEOP-LSTPLSHSTYVFLMWLFSLV 488  
DB 476 LA-ALVLL 483  
OY 489 FTVAITIGL 497

RESULT 9  
ID 075355 PRELIMINARY; PRT; 529 AA.  
AC 075355;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, last sequence update)  
DT 01-MAY-1999 (Tremblrel. 10, last annotation update)  
DE CD39L3.  
GN CD39L3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97419269.  
RA CHADWICK B.P., FRISCHAUF A.M.;  
RT "Cloning and mapping of a human and mouse gene with homology to ecto-  
RT ATPase genes.";  
RL Mamm. Genome 8:668-672(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 98341119.  
RA CHADWICK B.P., FRISCHAUF A.M.;  
RT "The CD39-like gene family: identification of three new human members  
RT the gene family from Drosophila melanogaster.";  
RL EMBL; AF039917; AAC39884.1; -  
DR PFAM; PF01150; GDAL\_CD39; 1.  
SQ SEQUENCE 529 AA; 59133 MW; 1AC2F1BA CRC32;

Query Match 35.3%; Score 1344; DB 4; Length 529;  
Best Local Similarity 38.9%; Pred. No. 6,60e-241;  
Matches 189; Conservative 115; Mismatches 172; Indels 10; Gaps 8;

DB 27 LVVL-LVSIIVLVSTIVQIHKEVLPRLGKYGIVLDAGSSRTTVYVQWPAEKENNTGV 85  
OY 18 LAIFGSSIIAVIALAVGLTKALPENVKYGIIVLDAGSSHTSLYIKWPAEKENDTG 77  
DB 86 VSGTFKSVKSGSISSYGNPNPVPRAFECKMVKGVPSHLHGSPHILGATAGMRL 145  
OY 78 VHOVEECRVKPGISKRVOKVNEIGIYLTDCMERAREVIRPSOHEPYYLGATAGMRL 137  
DB 146 RLONETANANVLESISQSYFKSOPDFRGAQIISGOEGVYGMTITANLGNLEKMLMH 205  
OY 138 RMESELDADVLDVERSLSNYPFDGARIITGOEGAGWTITINLLGKFSOKTRMFS 197  
DB 206 WV--HPRGVETGALDGGASTQISFVAGEKMDLNTSDIMOVSLYGVYTLVTHSFQCYG 263  
OY 198 IVPYETNNQTFGALDGGASTQVTFVPQNOTIESPDNALQFRLYGKDYNVVTHSFCLY 257  
DB 264 RNEAEKFLAMLQNSPTKHLNTPCYPRDYSSFTMGHVFSLCYTDQRPESINPDVI 323  
OY 258 KQDALMOKLAKDIOVA-SNEILRDPGFHPGKRVVNSDLYKTPCT--KRFEMTLRPOQF 314  
DB 324 TFEGTGDPSLCKEKVASIFPFKACHOETCSFQGVYOPKIKGPFVAFAGFYTASALNS 383  
OY 315 EIOGIGNYQOCHOSILELFTSYCPYSQ-CAFGIGIFLPLQGDGAFSAFYVWKFLNLT 373  
DB 384 GS-FSLDTFNSSWTNFCQSNWSQPLLLPKFDEVYARSYCFSANITYHLFVNGYKFTTEET 442  
OY 374 SEKVSGEKVTEMMKRFCAQPMWEIKTSYAGVKEKYLESCFSGTYILSLLOGVHTADS 433  
DB 443 WPOIHFEKVGNSSIAMSLGYMLSTNQIPAESPLRLPEPPVFGTIAFTTARLLCT 502  
OY 434 WEHIFHIGKIOGSDAGWTGLGMLNTNMIPAEOP-LSTPLSHSTYVFLMWLFSLV-LFTV 491

DB 503 AFLAYL 508  
OY 492 AITIGL 497

RESULT 10  
ID 060495 PRELIMINARY; PRT; 529 AA.  
AC 060495;  
DT 01-AUG-1998 (Tremblrel. 07, Created)  
DT 01-AUG-1998 (Tremblrel. 07, last sequence update)  
DT 01-MAY-1999 (Tremblrel. 10, last annotation update)  
DE E-TYPE ATPASE.  
GN HB6.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RA SMITH T.M., KIRLEY T.L.;  
RL Biochim. Biophys. Acta 0:0-0(1998).  
DR EMBL; AF034840; AAC09236.1; -  
DR PFAM; PF01150; GDAL\_CD39; 1.  
SQ SEQUENCE 529 AA; 59190 MW; 14BD39CD CRC32;

Query Match 35.2%; Score 1340; DB 4; Length 529;  
Best Local Similarity 38.9%; Pred. No. 4,47e-240;  
Matches 189; Conservative 114; Mismatches 173; Indels 10; Gaps 8;

DB 27 LVVL-LVSIIVLVSTIVQIHKEVLPRLGKYGIVLDAGSSRTTVYVQWPAEKENNTGV 85  
OY 18 LAIFGSSIIAVIALAVGLTKALPENVKYGIIVLDAGSSHTSLYIKWPAEKENDTG 77  
DB 86 VSGTFKSVKSGSISSYGNPNPVPRAFECKMVKGVPSHLHGSPHILGATAGMRL 145  
OY 78 VHOVEECRVKPGISKRVOKVNEIGIYLTDCMERAREVIRPSOHEPYYLGATAGMRL 137  
DB 146 RLONETANANVLESISQSYFKSOPDFRGAQIISGOEGVYGMTITANLGNLEKMLMH 205  
OY 138 RMESELDADVLDVERSLSNYPFDGARIITGOEGAGWTITINLLGKFSOKTRMFS 197  
DB 206 WV--HPRGVETGALDGGASTQISFVAGEKMDLNTSDIMOVSLYGVYTLVTHSFQCYG 263  
OY 198 IVPYETNNQTFGALDGGASTQVTFVPQNOTIESPDNALQFRLYGKDYNVVTHSFCLY 257  
DB 264 RNEAEKFLAMLQNSPTKHLNTPCYPRDYSSFTMGHVFSLCYTDQRPESINPDVI 323  
OY 258 KQDALMOKLAKDIOVA-SNEILRDPGFHPGKRVVNSDLYKTPCT--KRFEMTLRPOQF 314  
DB 324 TFEGTGDPSLCKEKVASIFPFKACHOETCSFQGVYOPKIKGPFVAFAGFYTASALNS 383  
OY 315 EIOGIGNYQOCHOSILELFTSYCPYSQ-CAFGIGIFLPLQGDGAFSAFYVWKFLNLT 373  
DB 384 GS-FSLDTFNSSWTNFCQSNWSQPLLLPKFDEVYARSYCFSANITYHLFVNGYKFTTEET 442  
OY 374 SEKVSGEKVTEMMKRFCAQPMWEIKTSYAGVKEKYLESCFSGTYILSLLOGVHTADS 433  
DB 443 WPOIHFEKVGNSSIAMSLGYMLSTNQIPAESPLRLPEPPVFGTIAFTTARLLCT 502  
OY 434 WEHIFHIGKIOGSDAGWTGLGMLNTNMIPAEOP-LSTPLSHSTYVFLMWLFSLV-LFTV 491  
DB 503 AFLAYL 508  
OY 492 AITIGL 497

RESULT 11  
ID 015464 PRELIMINARY; PRT; 472 AA.  
AC 015464;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, last sequence update)  
DT 01-NOV-1999 (Tremblrel. 12, last annotation update)  
DE ECTO-ATPASE (EC 3.6.1.5) (CD39 ANTIGEN-LIKE 1).

GN CD39L1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 97419269.  
RA CHADWICK B.P., FRISCHAUF A.-M.;  
RT "Cloning and mapping of a human and mouse gene with homology to ecto-  
ATPase genes."  
RL Mamm. genome 8:668-672(1997).  
CC -1- FUNCTION: HYDROLYSES EXTRACELLULAR ATP TO AMP.  
CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -1- TISSUE SPECIFICITY: BRAIN, PLACENTA, SKELETAL MUSCLE, KIDNEY,  
PANCREAS, HEKT, OVARY, TESTIS, COLON, SMALL INTESTINE, PROSTATE,  
AND PANCREAS.  
CC -1- PPM: HAS PROBABLY VARIOUS DISULFIDE-BONDS (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.  
DR EMBL: U91510; AAB81013.1; -.  
DR MIM: 602012; -.  
DR PROSITE: PS01238; GDAL-CD39\_NTPASE; 1.  
DR PFM: PFM01150; GDAL-CD39; 1.  
KW Hydrolyase; Transmembrane; Antigen; Glycoprotein.  
FT DOMAIN 1 7  
FT TRANSSEM 8 28  
FT DOMAIN 22 25  
FT DOMAIN 29 439  
FT TRANSSEM 440 460  
FT DOMAIN 444 447  
FT DOMAIN 461 472  
FT CARBOHYD 64 64  
FT CARBOHYD 129 129  
FT CARBOHYD 294 294  
FT CARBOHYD 378 378  
FT CARBOHYD 420 420  
FT SEQUENCE 472 AA; 51161 MM; 5A924C38 CRC32;  
Query Match 29.3%; Score 1115; DB 4; Length 472;  
Best Local Similarity 40.5%; Pred. No. 1.58e-193;  
Matches 197; Conservative 111; Mismatches 143; Indels 36; Gaps 24;

DB 455 -ALVILL 460  
QY 491 VALIGEL 497  
RESULT 12  
ID 09X162 PRELIMINARY; PRT: 483 AA.  
AC 09X162;  
DT 01-NOV-1999 (Tremblrel. 12, Created)  
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)  
DE F7A19.34 PROTEIN.  
GN F7A19.34.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC eudicotyledons; Spermatophytes; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CD, COLUMBIA;  
RA FEEDERSTIEL N.A., PALM C.J., CONWAY A.B., CONN L., HANSEN N.F.,  
RA ALFAEI H., ARAUJO R., HUIZAR L., ROWLEY D., BUEHLER E., DUNN P.,  
RA GONZALEZ A., KREMENTSKAYA I., KIM C., LENZ C., LI J., LIU S.,  
RA LUKOS S., SCHWARTZ J., SHINN P., TORIUMI M., VYSOTSKAIA V.S.,  
RA WALKER M., YU G., ECKER J., THEOLOGIS A., DAVIS R.W.;  
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AC007576; AAD39311.1; -.  
DR SEQUENCE 483 AA; 53425 MM; B134313C CRC32;  
Query Match 15.5%; Score 590; DB 10; Length 483;  
Best Local Similarity 30.9%; Pred. No. 4.59e-87;  
Matches 131; Conservative 104; Mismatches 154; Indels 35; Gaps 25;

DB 67 KLRVSLIDASSGTRVAVFGVWFSGKVPDGEKHYANL-KLT-PGLSSYADNPEGAS 124  
QY 46 NKIKIVLDAGSSSHSLIYIK-W-PAEKEN-DTGCVHGYEBCRVGPGISFVQVANEIG 102  
DB 125 VSVTKLVEFAKORLPKRMFRSDIRIMATAGKRL-EVP-VOEQLLETRVRLSSGFM 181  
QY 103 IYLPDCMERAREVIRPSQHOENPVYLAGTAGKRLRMSEBELADVLVVERSLSNYPED 162  
DB 182 FRDEMANVSSDGCISYITANTALGSLG--T--D--PLET-----T-GIVEAGASAO 229  
QY 163 FOG--ARIITQEEGAYGMITINILGKFSQKTRMFSTVPEYETNQEFFGALDLAGASTQ 220  
DB 230 VTFVSSSEH-V-PPESRTIA-YGNISYITISHSFLDYGKDAL-KLLEKIONSANSTV- 284  
QY 221 VTFVPONOTIESPDNALQFRLYGR-DYVNVTHSEFLCYGKDQALQKLDQVANSNETLR 279  
DB 285 DGVEDPCTPGKYIDYTRSKNYSGLFLADESKLKSLOAANFSKRSRATFALLKEGEN 344  
QY 280 DPCFHPGKKVYVNSDLKTPCTKRFEMTLRPOQEEIGICIGNYQCHOSILEPFTS--Y 337  
DB 345 CLYEHCSIGSTFTPDGSLATASFTYAKFPEL-BEKGWLSLIPAGKRYGGEWSKL 403  
QY 338 CPYSCAFNGIFLPLQDGFCAFSAFYVMKFLINTSEKVSQEKTEMKKFCAPWBEI 397  
DB 404 ILEVPTDEEYLRGCFSAATYSMLHDSLGIALDD-SITYASAGAKGKHIPDLQALAF 462  
QY 398 KTSYAGVEKYLSEICFSGITILSLDLOGYHFTADSWEHIFHIGIKQSD-A-GWTLG-Y 454  
DB 463 ILDV 466  
QY 455 MLNL 458  
RESULT 13  
ID 015092 PRELIMINARY; PRT: 609 AA.  
AC 015092;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)



```

OY 17 IIAIIIGFSSIIIAVIAL-LAVGLTQNKALPENVKYGIULDAGSSHTSLYTKMPAEKENDT 75
DB 95 FEFRGANVYASLKLHPGLSAPADPOGASVSLTELVEFAKGRVPGKMWIETEVRLMATAGM 154
OY 76 GYVHQVEECRRK-GPGISKFEVOKVNEIGIYLTDCMERAREVILPRSOHQETPVYLGATAGM 134
DB 155 RLLEL-PVQ-EKILGVARRVLKSSGFLPRDEMASVYSGSDGVYAWVYANFALGSLG-- 209
OY 135 RLIRMESELADRVLDVVERSLSNYPFDQ--ARITIQEGAGYGMITINYLKFSQK 192
DB 210 G--D--PLKT---T-GIVEIGASAOVTVS-SEPM-PPEPSRTIS-FGNVTYNLYSH 256
OY 193 TWMFSTIVPYETNNQETFGALDYGASTQYTFVPONOTIESPDNALQFRLYK-DYVNYTH 251
DB 257 SFLHFGQNAH-DKLMGSL-L-SRD--HNSAVEPT-RE-KT--FTDPCARKYNLDAN 305
OY 252 SFLCYGKDDALMOKLAKDIQVASNELTRDPCFHPGKYKVVNVSPLYKTPCT-KREMTLP 310
DB 306 TQK-HLSGLA-EESRLS--DSFOAG-GNYSOCRSALITLQDNGRILITLAGFSF-L- 358
OY 311 PQOFELQIGINYQOCHQSTILELFTNTSYCPYSQCAFNGI-FLPPLQGF-GAFSAFYVMK 368
DB 359 FFGI-GEKAMLSNMTSAGERFCGEDMSKLRVDPSLHEEDLRYCFSSAYIVSLDHTLG 417
OY 369 FLNLTSEKYSQEKYTEMMKKFCAQPMEEIKTSYAGVKEKYLSEYCFSGTYILSLQGYH 428
DB 418 IPLDD-ERIGYANQAGDIPLDALGAFIOQTATETSQHNASGNL 460
OY 429 FTADSWEHIFIGKIQSDAGMTLGYMLNTNMIPAEQPLSTPL 472

```

Search completed: Fri May 5 08:58:53 2000  
 Job time : 504 secs.





QY 121 NYPDFOGARIITGOEGAGWITTYINVLGKFSQKTRMFSIVPYETNNQETFGALDLGGA 180  
 DB 218 STQVTEVPQNOTIESPDNALQFRLYKGDVNYTHSFLCYCKDQALMOKLAKDIOVASNEI 277  
 QY 181 STQVTEVPQNOTIESPDNALQFRLYKGDVNYTHSFLCYCKDQALMOKLAKDIOVASNEI 240  
 DB 278 LRDPCEHPGKRVKVVNSDLYKTPCTKRFEMTLPPQOFEIQIGNYQCHQSILEFNTSY 337  
 QY 241 LRDPCEHPGKRVKVVNSDLYKTPCTKRFEMTLPPQOFEIQIGNYQCHQSILEFNTSY 300  
 DB 338 CPYSQCAFNGIFLPLPLOGDFGASAFYFVKFLNLTSEKVSQEKVTEMMKKFCQAPWEEI 397  
 QY 301 CPYSQCAFNGIFLPLPLOGDFGASAFYFVKFLNLTSEKVSQEKVTEMMKKFCQAPWEEI 360  
 DB 398 KTSYAGVKEKYISEYCSGTYIISLLQGYHFTADSWEHIFHGKIQGSDAGWTLGYMLN 457  
 QY 361 KTSYAGVKEKYISEYCSGTYIISLLQGYHFTADSWEHIFHGKIQGSDAGWTLGYMLN 420  
 DB 458 LTNMIRPAEQPLSTPLSHST 476  
 QY 421 LTNMIRPAEQPLSTPLSHST 439

## RESULT 2

ID W04264 standard; Protein: 510 AA.  
 AC W04264;  
 DT 29-MAY-1997 (first entry)  
 DE Human CD39 protein.  
 KW Human; lymphocyte activation marker; gene therapy;  
 KW ATP diphosphohydrolase; transgenic animal; pig; porcine; graft;  
 KW donor; transplantation; endothelial cell; prosthetic device;  
 KW platelet aggregation; inhibition; intravascular.  
 OS Homo sapiens.  
 PN W09630532-A1.  
 PD 03-OCT-1996.  
 PR 22-MAR-1996; E01270.  
 PR 24-MAR-1995; US-410371.  
 PR 12-FEB-1996; US-600383.  
 PA (NEMO-) NEW ENGLAND DEACONESS HOSPITAL.  
 PA (SANO) SANDOZ LTD.  
 PI Bach FH, Rodson S;  
 DR WPI: 96-455377/45.  
 DR N-PEDB: T33966.  
 PT Gene therapy of inflammatory or immunological stimulation of  
 PT platelet aggregation - using CD39 protein with ATP  
 PT di:phospho:hydrolase activity, useful for preventing or alleviating  
 PT thrombotic condition in mammalian subject  
 PS Claim 3; Page 39; 65pp; English.  
 CC Non-human transgenic or somatic recombinant mammals, whose cells  
 CC contain a heterologous DNA encoding a polypeptide (especially human  
 CC CD39 protein) having ATP-diphosphohydrolase activity under cellular  
 CC activating conditions is claimed. In particular the animal is a pig  
 CC and its cells (or tissues or organs) can be used for transplantation.  
 CC DNA coding for human CD39 is also useful for genetically modifying a  
 CC mammalian cell to render it less susceptible to an inflammatory or  
 CC immunological stimulus and platelet aggregation. The modified cells  
 CC can be used to prevent or alleviate a thrombotic condition.  
 CC The present sequence is that of the human CD39 protein and was  
 CC disclosed in J. Immunol. 153 (8) (1994) 3574-3584.  
 SQ Sequence 510 AA;

Query Match 100.0%; Score 3275; DB 1; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
 Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 38 TONKALPENVKYGIYVADAGSSHTSLYIKWPAKENDTGVVHVECRVKGPGISKRVOK 97  
 QY 1 TONKALPENVKYGIYVADAGSSHTSLYIKWPAKENDTGVVHVECRVKGPGISKRVOK 60  
 DB 98 VNEIGIYLTDCHEARAREVIPRSOHOETPYVLGATAGMRLIMSEELADRVLDYVERSLS 157  
 QY 61 VNEIGIYLTDCHEARAREVIPRSOHOETPYVLGATAGMRLIMSEELADRVLDYVERSLS 120

DB 158 NYPDFOGARIITGOEGAGWITTYINVLGKFSQKTRMFSIVPYETNNQETFGALDLGGA 217  
 QY 121 NYPDFOGARIITGOEGAGWITTYINVLGKFSQKTRMFSIVPYETNNQETFGALDLGGA 180  
 DB 218 STQVTEVPQNOTIESPDNALQFRLYKGDVNYTHSFLCYCKDQALMOKLAKDIOVASNEI 277  
 QY 181 STQVTEVPQNOTIESPDNALQFRLYKGDVNYTHSFLCYCKDQALMOKLAKDIOVASNEI 240  
 DB 278 LRDPCEHPGKRVKVVNSDLYKTPCTKRFEMTLPPQOFEIQIGNYQCHQSILEFNTSY 337  
 QY 241 LRDPCEHPGKRVKVVNSDLYKTPCTKRFEMTLPPQOFEIQIGNYQCHQSILEFNTSY 300  
 DB 338 CPYSQCAFNGIFLPLPLOGDFGASAFYFVKFLNLTSEKVSQEKVTEMMKKFCQAPWEEI 397  
 QY 301 CPYSQCAFNGIFLPLPLOGDFGASAFYFVKFLNLTSEKVSQEKVTEMMKKFCQAPWEEI 360  
 DB 398 KTSYAGVKEKYISEYCSGTYIISLLQGYHFTADSWEHIFHGKIQGSDAGWTLGYMLN 457  
 QY 361 KTSYAGVKEKYISEYCSGTYIISLLQGYHFTADSWEHIFHGKIQGSDAGWTLGYMLN 420  
 DB 458 LTNMIRPAEQPLSTPLSHST 476  
 QY 421 LTNMIRPAEQPLSTPLSHST 439

## RESULT 3

ID W04339 standard; Protein: 20 AA.  
 AC W04339;  
 DT 29-DEC-1996 (first entry)  
 DE ATP diphosphohydrolase fragment from pig pancreas.  
 KW ATP diphosphohydrolase; ATPase; bovine aorta; pig pancreas;  
 KW apyrase; CD39; lymphoid cell activation antigen; enzyme;  
 KW platelet aggregation; thrombogenicity; anti-haemostatic.  
 OS Sus scrofa.  
 PN W09632471-A2.  
 PD 17-OCT-1996.  
 PR 10-APR-1996; CA0223.  
 PR 10-APR-1995; US-419204.  
 PA (UYSH ) UNIV SHERBROOKE.  
 PI Beaudoin AR, Sevigny J;  
 DR WPI: 96-477122/47.  
 DR Isolated ATP di:phospho:hydrolase enzymes - have anti-haemostatic  
 PT activity, useful for reducing platelet aggregation and  
 PT thrombogenicity  
 PS Claim 3; Page 46; 60pp; English.  
 CC The bovine aorta (W04339) ATPases have been partially sequenced.  
 CC Pancreatic (W04339) ATPases have been partially sequenced.  
 CC The sequences have been found to be highly homologous to a human  
 CC lymphoid cell activation antigen designated CD39 (Maliszewski et  
 CC al. (1994). J. Immunol.: 3574-3583). The complete sequences of the  
 CC ATPases types I and II have not been obtained yet. Assuming that  
 CC the CD39 gene product is an ATPase type II, the use of CD39  
 CC in the reduction of platelet aggregation and of thrombogenicity  
 CC may be contemplated, as well as a process of making ATPases  
 CC using the CD39 sequence (T38516).  
 SQ Sequence 20 AA;

Query Match 3.2%; Score 104; DB 1; Length 20;  
 Best Local Similarity 78.9%; Pred. No. 1.03e+01;  
 Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 2 SDQETFGALDLGASTOV 20  
 QY 166 TNNQETFGALDLGASTOV 184

## RESULT 4

ID W20122 standard; Protein: 248 AA.  
 AC W20122;  
 DT 07-JUL-1997 (first entry)  
 DE H. pylori surface membrane protein, 12969218.aa.  
 KW Cytoplasmic; vaccine; prevention; infection; identification;  
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;



FT	misc_difference	269	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	285	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	312	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	319	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	344	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	358	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	378	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	386	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	465	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	473	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	473	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	494	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	555	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	593	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	602	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	609	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	621	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	724	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	736	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	739	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	786	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	841	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	924	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	934	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	1017	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	1054	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	1127	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	1147	/note=	"corresponds to stop codon in DNA sequence"
FT	misc_difference	1178	/note=	"corresponds to stop codon in DNA sequence"
PN	W09632476-A1.			
PD	17-OCT-1996.			
PF	12-APR-1996; CA0214.			
PR	13-APR-1995; US-421701.			
PA	(MOUN ) MOUNT SINAI HOSPITAL CORP.			
PI	Labes M, Lozano A, Roach A, Roder J;			
DR	WPI; 96-477127/47.			
DR	N-PSDB; T38484.			
PT	Assay for substance that modulates response of neuronal cells - and			
PT	neurite growth associated protein, Petrin, useful in conditions			
PT	involving nerve damage resulting from traumatic injury, stroke or			
PT	CNS degenerative disorders			
PS	Claim 9; Page 57-61; 119pp; English.			
CC	Rat petrin (W04326) is a protein involved in modulating neurite			
CC	growth inhibition. The amino sequence was deduced from a cDNA			

CC clone (T38484) derived from an adult rat brain cDNA library; no  
 CC coding sequence was indicated. Petrin is a new member of the  
 CC protein phosphatase 2C family, and is expressed in neurons in brain  
 CC tissue, partic. in the Purkinje cells of the cerebellum. Petrin,  
 CC and antibodies raised against it, can be used to modulate neurite  
 CC growth and axonal regeneration.  
 SQ Sequence 1196 AA;

Query Match 3.2%; Score 106; DB 1; Length 1196;  
 Best Local Similarity 31.4%; Pred. No. 7,33e+00;  
 Matches 16; Conservative 15; Mismatches 18; Indels 2; Gaps 2;

Db 503 KYEGSHCVQLFQSGFLPISCPLEFISFCGDSGVRSFQGFVFL-FL 552  
 OY 284 NYQCHQSHLIEFNTSYCPYSQAFN-GIFLPLQDGFCAFSAFYVMFL 333

RESULT 6  
 ID W20877; standard; Protein; 95 AA.

DE 18-JUL-1997 (first entry)  
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacterium; life cycle; activator;  
 KW bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;  
 KW diagnosis.

OS Helicobacter pylori.

PN M09640893-A1.

PD 19-DEC-1996.

PF 06-JUN-1996; U09122.

PR 07-JUN-1995; US-487032.

PR 01-APR-1996; US-630405.

PA (ASTR) ASTRA AB.

PI Berglindh OT, Smith D, Mellegaard BL;

DR WPI: 97-052306/05.

DR N-PSDB: T68130.

PT Helicobacter pylori nucleic acid sequences and related  
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
 PT infection, and to detect Helicobacter

PS Claim 61; Page 1278-79; 1481pp; English.

CC This sequence represents a H. pylori cytoplasmic protein.

CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds.

CC useful as potential H. pylori life cycle activators or inhibitors.

CC The genomic sequence of H. pylori (ATCC 55679) was determined from  
 CC overlapping contigs generated by mechanically shearing the bacterial  
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,

CC and the predicted coding regions defined by computer evaluation. To  
 CC identify likely H. pylori antigens for vaccine development, the amino  
 CC acid sequences predicted from various ORF were analysed for significant

CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide

CC production, e.g. in E. coli hosts.

SQ Sequence 95 AA;

Query Match 3.1%; Score 103; DB 1; Length 95;  
 Best Local Similarity 34.0%; Pred. No. 1,21e+01;  
 Matches 18; Conservative 16; Mismatches 14; Indels 5; Gaps 5;

Db 24 GYM-VSLGALKRPTCTNRFYKALLFAIF-YHAVNNFLTQCPHGVREFFSS 74  
 OY 249 GYKKVNVSDLYKPTCKRFEM-TLPFOQFELQIGNTY-QQCH-GSILBLFMT 298

RESULT 7  
 ID R85880; standard; Protein; 439 AA.

AC R85880;

DE 13-SEP-1996 (first entry)

KW WD-40 domain-contg. YCUT7 protein.  
 KW WD40 repeat region; beta-transducin; protein-protein interaction; drug;  
 KW intracellular signalling; protein kinase C; homology; motif; modulator;  
 KW receptors of activated protein kinase; enzyme activity; isozyme; human.

OS Synthetic.

PN W09521252-A2.

PD 10-AUG-1995.

PF 01-JAN-1995; U01210.

PR 01-FEB-1994; US-190802.

PA (STND) UNIT LELAND STANFORD JUNIOR.

PI Mochly-Rosen D, Ron D;

DR WPI: 95-283772/37.

PT New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the  
 PT activity of a protein, eg. protein kinase C, which interacts with a  
 PT protein contg. a WD-40 region.

PS Example 5; Page 166-168; 351pp; English.

CC Proteins R85851-92 are protein in which contain at least one WD-40 (also  
 CC called beta-transducin homologous) amino acid repeat motifs. The WD-40  
 CC regions are involved in protein-protein interactions between proteins

CC involved in intracellular signalling. An example of such an interaction  
 CC is between protein kinase C and receptors of activated protein kinase  
 CC (RACK), esp. RACK-1 (R85850). Proteins R85851-82 were isolated based on  
 CC homology with beta-transducin, whereas proteins R85882-92 were isolated  
 CC based on homology with the WD-40 consensus sequence (R85893). The

CC R85786-R85842. The peptides can be used to identify target proteins  
 CC contg. WD-40 motifs, as modulators of enzyme esp. isozyme, activity of  
 CC proteins involved in protein-protein interaction and to screen for drugs  
 CC that will affect protein-protein interaction involving WD-40 domains.

SQ Sequence 439 AA;

Query Match 3.1%; Score 102; DB 1; Length 439;  
 Best Local Similarity 26.7%; Pred. No. 1,43e+01;  
 Matches 16; Conservative 16; Mismatches 23; Indels 3; Gaps 3;

Db 180 AYSPPDSRVYVASEDKIKVMDITSGFCLATEEHTSSVTAVQFARKQVEMSS-SLDGT 238  
 OY 123 PFDQGARITITGEGAY-GW-ITINYLKRSOKTRMWSIVPYETNNQDFGALDGA 180

RESULT 8  
 ID W20482; standard; Protein; 70 AA.

AC W20482;

DE 15-JUL-1997 (first entry)

KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.

OS Helicobacter pylori.

PN M09640893-A1.

PD 19-DEC-1996.

PF 06-JUN-1996; U09122.

PR 07-JUN-1995; US-487032.

PR 01-APR-1996; US-630405.

PA (ASTR) ASTRA AB.

PI Berglindh OT, Smith D, Mellegaard BL;

DR WPI: 97-052306/05.

DR N-PSDB: T67645.

PT Helicobacter pylori nucleic acid sequences and related  
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
 PT infection, and to detect Helicobacter

PS Claim 61; Page 648-649; 1481pp; English.

CC The present sequence shows a Helicobacter pylori cytoplasmic protein  
 CC that may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,

CC useful as potential H. pylori life cycle activators or inhibitors.

CC The genomic sequence of H. pylori (ATCC 55679) was determined from  
 CC overlapping contigs generated by mechanically shearing the bacterial  
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,

CC and the predicted coding regions defined by computer evaluation. To  
 CC identify likely H. pylori antigens for vaccine development, the amino  
 CC acid sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts.  
 SQ Sequence 70 AA;



PA (VUSL-) UNIV ST LOUIS.  
PI Chinnadurai G;  
DR WPI: 99-152099/13.  
PT Polypeptides that bind to anti-apoptotic proteins - useful for

PT protecting against cell death induced by viral infection and to  
PR modulate response to physical and chemical stimuli  
PS Example 6: Column 39-42: 41pp. english  
CC The present invention describes: (1) a method for regulating cell death,  
CC comprising exposing an isolated cell to a polypeptide selected from  
CC N1p1, N1p2, N1p3, B1p1, B1p2 and B1p3; (2) a method for neutralising  
CC the activity of the adenovirus E1B 19 kD protein, the Bcl-2 protein or  
CC the BHRF-1 protein, comprising exposing an isolated cell to a  
CC polypeptide as in (1); and (3) a method for detecting molecules that  
CC bind to at least one polypeptide as in (1), comprising lysing cells,  
CC exposing the lysate to the polypeptide and detecting any molecule-  
CC polypeptide aggregates. The methods are useful for providing proteins  
CC able to bind to other proteins known to regulate cell survival e.g. it  
CC is known that E1B 19k protein provides a survival function similar to  
CC the cellular protooncogene bcl-2 gene product which is able to block  
CC apoptosis in haematopoietic B and T cells. The present sequence  
CC represents an adenovirus mutant 19k protein from the present invention.  
SQ Sequence 147 AA:

Query Match	2.88;	Score 93;	DB 1;	Length 147;
PostgreSQL	30.18	93	1	147
PostgreSQL	30.18	93	1	147

Best Local Similarity 29.18; Pred. No. 6.22e+01;  
Watched 15; Generated 12; Wasted 34

Matches	16;	Conservative	13;	Mismatches	24;	Indels	2;	Gaps	2;
---------	-----	--------------	-----	------------	-----	--------	----	------	----

Db 65 LNLGHQALFQEKVIKTLD-FSTPGRAAAVAFLSFINKDWSEETHLSGGYLLDFL 118

0Y 333 LNLTSKVSQEKYJENMKKFCQAQPMEEIKTSYAG-VKEKYLSEYCFSGFYILSL 386

Search completed: Fri May 5 09:18:10 2000  
Job time : 157 secs.



\*\*\*\*\*  
M P E S E R E H  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri May 5 09:35:11 2000; MasPar time 13.74 Seconds

Tabular output not generated. 414.072 Million cell updates/sec

Title: >US-09-374-586-2  
Description: (1-439) from US09374586.pep  
Perfect Score: 3275  
Sequence: 1 TONKALPENVKYGVLDAGS.....NLTNMPAEPFLSTPLSHST 439

Scoring table: PAM 150  
Gap 11

Searched: 131253 segs, 12956647 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-Issued  
1:5A\_COMB 2:5B\_COMB 3:PC9\_COMB 4:backfiles1

Statistics: Mean 33.222; Variance 147.681; scale 0.225

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	102	3.1	439	1 US-08-190-	Sequence 65, Applicati	7.00e+00
2	96	2.9	319	1 US-08-413-	Sequence 119, Applicati	1.86e+01
3	96	2.9	459	1 US-08-220-	Sequence 12, Applicati	1.86e+01
4	96	2.9	459	1 US-08-220-	Sequence 14, Applicati	1.86e+01
5	96	2.9	459	1 US-08-413-	Sequence 12, Applicati	1.86e+01
6	96	2.9	459	1 US-08-413-	Sequence 14, Applicati	1.86e+01
7	95	2.9	1285	1 US-07-582-	Sequence 2, Applicatio	2.19e+01
8	95	2.9	1285	1 US-08-453-	Sequence 2, Applicatio	2.19e+01
9	93	2.8	147	2 US-08-408-	Sequence 15, Applicati	3.02e+01
10	93	2.8	175	2 US-08-408-	Sequence 14, Applicati	3.02e+01
11	93	2.8	175	2 US-08-408-	Sequence 11, Applicati	3.02e+01
12	93	2.8	175	2 US-08-408-	Sequence 10, Applicati	3.02e+01
13	93	2.8	175	2 US-08-408-	Sequence 9, Applicatio	3.02e+01
14	93	2.8	175	2 US-08-408-	Sequence 8, Applicatio	3.02e+01
15	92	2.8	1025	2 US-08-530-	Sequence 23, Applicati	3.54e+01
16	90	2.7	205	4 5175383-6	Patent No. 5175383.	4.85e+01
17	88	2.7	337	2 US-08-806-	Sequence 2, Applicatio	6.64e+01
18	88	2.7	352	2 US-08-751-	Sequence 2, Applicatio	6.64e+01
19	88	2.7	361	2 US-08-415-	Sequence 36, Applicati	6.64e+01
20	87	2.7	440	2 US-08-485-	Sequence 31, Applicati	7.75e+01
21	87	2.7	440	1 US-08-483-	Sequence 27, Applicati	7.75e+01
22	87	2.7	474	1 US-08-459-	Sequence 4, Applicatio	7.75e+01
23	87	2.7	678	1 US-08-288-	Sequence 5, Applicatio	7.75e+01

24	87	2.7	1089	1 US-08-180-	Sequence 36, Applicati	7.75e+01
25	87	2.7 <th>1089</th> <th>2 US-08-475-</th> <th>Sequence 4, Applicati</th> <th>7.75e+01</th>	1089	2 US-08-475-	Sequence 4, Applicati	7.75e+01
26	87	2.7 <th>1089</th> <th>1 US-08-460-</th> <th>Sequence 4, Applicati</th> <th>7.75e+01</th>	1089	1 US-08-460-	Sequence 4, Applicati	7.75e+01
27	87	2.7 <th>1089</th> <th>1 US-08-168-</th> <th>Sequence 4, Applicati</th> <th>7.75e+01</th>	1089	1 US-08-168-	Sequence 4, Applicati	7.75e+01
28	87	2.7 <th>1089</th> <th>3 PCT-US92-0</th> <th>Sequence 4, Applicati</th> <th>7.75e+01</th>	1089	3 PCT-US92-0	Sequence 4, Applicati	7.75e+01
29	87	2.7 <th>1089</th> <th>2 PCT-US92-0</th> <th>Sequence 4, Applicati</th> <th>7.75e+01</th>	1089	2 PCT-US92-0	Sequence 4, Applicati	7.75e+01
30	87	2.7 <th>1089</th> <th>2 US-08-460-</th> <th>Sequence 4, Applicati</th> <th>7.75e+01</th>	1089	2 US-08-460-	Sequence 4, Applicati	7.75e+01
31	87	2.7 <th>1089</th> <th>1 US-08-477-</th> <th>Sequence 36, Applicati</th> <th>7.75e+01</th>	1089	1 US-08-477-	Sequence 36, Applicati	7.75e+01
32	89	2.7 <th>2182</th> <th>2 US-08-487-</th> <th>Sequence 16, Applicati</th> <th>5.68e+01</th>	2182	2 US-08-487-	Sequence 16, Applicati	5.68e+01
33	85	2.6 <th>163</th> <th>1 US-08-344-</th> <th>Sequence 7, Applicatio</th> <th>1.06e+02</th>	163	1 US-08-344-	Sequence 7, Applicatio	1.06e+02
34	85	2.6 <th>163</th> <th>1 US-07-991-</th> <th>Sequence 7, Applicatio</th> <th>1.06e+02</th>	163	1 US-07-991-	Sequence 7, Applicatio	1.06e+02
35	85	2.6 <th>163</th> <th>1 US-08-107-</th> <th>Sequence 7, Applicatio</th> <th>1.06e+02</th>	163	1 US-08-107-	Sequence 7, Applicatio	1.06e+02
36	84	2.6 <th>263</th> <th>3 PCT-US94-0</th> <th>Sequence 2, Applicatio</th> <th>1.06e+02</th>	263	3 PCT-US94-0	Sequence 2, Applicatio	1.06e+02
37	86	2.6 <th>389</th> <th>1 US-07-939-</th> <th>Sequence 1, Applicatio</th> <th>9.05e+01</th>	389	1 US-07-939-	Sequence 1, Applicatio	9.05e+01
38	85	2.6 <th>417</th> <th>1 US-08-351-</th> <th>Sequence 6, Applicatio</th> <th>1.06e+02</th>	417	1 US-08-351-	Sequence 6, Applicatio	1.06e+02
39	85	2.6 <th>417</th> <th>1 US-08-351-</th> <th>Sequence 7, Applicatio</th> <th>1.06e+02</th>	417	1 US-08-351-	Sequence 7, Applicatio	1.06e+02
40	86	2.6 <th>423</th> <th>1 US-07-939-</th> <th>Sequence 10, Applicati</th> <th>9.05e+01</th>	423	1 US-07-939-	Sequence 10, Applicati	9.05e+01
41	86	2.6 <th>423</th> <th>1 US-07-939-</th> <th>Sequence 12, Applicati</th> <th>9.05e+01</th>	423	1 US-07-939-	Sequence 12, Applicati	9.05e+01
42	85	2.6 <th>617</th> <th>1 US-08-137-</th> <th>Sequence 25, Applicati</th> <th>1.06e+02</th>	617	1 US-08-137-	Sequence 25, Applicati	1.06e+02
43	84	2.6 <th>671</th> <th>4 5266464-2</th> <th>Patent No. 5266464.</th> <th>1.23e+02</th>	671	4 5266464-2	Patent No. 5266464.	1.23e+02
44	86	2.6 <th>1658</th> <th>2 US-08-609-</th> <th>Sequence 13, Applicati</th> <th>9.05e+01</th>	1658	2 US-08-609-	Sequence 13, Applicati	9.05e+01
45	86	2.6 <th>1726</th> <th>2 US-08-609-</th> <th>Sequence 30, Applicati</th> <th>9.05e+01</th>	1726	2 US-08-609-	Sequence 30, Applicati	9.05e+01

## ALIGNMENTS

RESULT	1	STANDARD;	PRT;	439 AA.
ID	US-08-190-802A-65			
XX	xxxxxx			
XX				
DE	Sequence 65, Application US/08190802A			
XX	Sequence 65, Application US/08190802A			
CC	Patent No. 5519003			
CC	GENERAL INFORMATION:			
CC	APPLICANT: Mochly-Rosen, Daria			
CC	APPLICANT: Ron, Dorit			
CC	TITLE OF INVENTION: WD-40 - Derived Peptides and Uses			
CC	TITLE OF INVENTION: Thereof			
CC	NUMBER OF SEQUENCES: 265			
CC	CORRESPONDENCE ADDRESS:			
CC	ADDRESSEE: Denlinger & Associates			
CC	STREET: P.O. Box 60850			
CC	CITY: Palo Alto			
CC	STATE: CA			
CC	COUNTRY: USA			
CC	ZIP: 94306-0850			
CC	COMPUTER READABLE FORM:			
CC	MEDIUM TYPE: Floppy disk			
CC	COMPUTER: IBM PC compatible			
CC	OPERATING SYSTEM: PC-DOS/MS-DOS			
CC	SOFTWARE: Patentin Release #1.0, Version #1.25			
CC	CURRENT APPLICATION DATA:			
CC	APPLICATION NUMBER: US/08/190,802A			
CC	CLASSIFICATION: 530			
CC	FILING DATE: 01-FEB-1994			
CC	ATTORNEY/AGENT INFORMATION:			
CC	NAME: Fabian, Gary R.			
CC	REGISTRATION NUMBER: 33,875			
CC	REFERENCE/DOCKET NUMBER: 8600-0139			
CC	TELECOMMUNICATION INFORMATION:			
CC	TELEPHONE: (415) 324-0880			
CC	TELEFAX: (415) 324-0960			
CC	INFORMATION FOR SEQ ID NO: 65:			
CC	SEQUENCE CHARACTERISTICS:			
CC	LENGTH: 439 amino acids			
CC	TYPE: amino acid			
CC	TOPOLOGY: unknown			
CC	MOLECULE TYPE: Protein			
CC	HYPOTHETICAL: NO			







```

CC      MOLECULE TYPE: protein
SO      SEQUENCE 1285 AA; 146565 MW; 8485412 CN;

Query Match      2.9%; Score 95; DB 1; Length 1285;
Best Local Similarity 30.6%; Pred. No. 2.19e+01;
Matches 19; Conservative 13; Mismatches 26; Indels 4; Gaps 4;

Db      823 PVELXPTFFFSLQVQSDLG-FEDAFATRFPEFTLVSDRLSLMENTMLTLESFDYTPMD 881
OY      302 PPSQCAFMGI-FLPPLQSDGFSAFYVMKFLN-LTSEKVS-OEKYTEMKKFCQAPWE 358
Db      882 AI 883
OY      359 EI 360

RESULT 8 STANDARD; PRT; 1285 AA.
ID      US-08-453-141-2
XX      AC      xxxxxx
XX      DT
XX      XX
XX      XX
XX      Sequence 2, Application US/08453141
DE      CC      Patent No. 5885589
XX      CC      GENERAL INFORMATION:
CC      APPLICANT: FOGED, Niels T.
CC      APPLICANT: PETERSEN, Svend
CC      TITLE OF INVENTION: PASTEDURELLA VACCINE
CC      NUMBER OF SEQUENCES: 2
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Foley & Lardner
CC      STREET: 3000 K Street, N.W., Suite 500
CC      CITY: Washington, D.C.
CC      COUNTRY: USA
CC      ZIP: 20007-5109
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/453,141
CC      FILING DATE:
CC      CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/293,314
CC      FILING DATE: 22-AUG-1994
CC      APPLICATION NUMBER: US 07/582,945
CC      FILING DATE: 12-OCT-1990
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: PCT/DK89/00084
CC      FILING DATE: 11-APR-1989
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: DK 1995/88
CC      FILING DATE: 04-DEC-1988
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: BENT, Stephen A.
CC      REGISTRATION NUMBER: 29,768
CC      REFERENCE/DOCKET NUMBER: 40399/102/AKZO
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (202)672-5300
CC      TELEFAX: (202)672-5399
CC      TELEX: 904136
CC      INFORMATION FOR SEQ ID NO: 2:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 1285 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 1285 AA; 146552 MW; 8486546 CN;

```



```
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/408,095  
CC FILING DATE: 21-MAR-1995  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Mack, Susan J.  
CC REGISTRATION NUMBER: 30,951  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (202)293-7060  
CC FAX: (202)293-7060  
CC TELEFAX: (202)293-7860  
CC INFORMATION FOR SEQ ID NO: 8:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 175 amino acids  
CC STRANDEDNESS:  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 175 AA; 20514 MW; 143104 CN;  
  
Query Match 2.8%; Score 93; DB 2; Length 175;  
Best Local Similarity 29.1%; Pred.No.3,02e+01;  
Matches 16; Conservative 13; Mismatches 24; Indels 2; Gaps 2  
  
Db 65 LNLGQALFOEKVITKLD-FSTPGRAAAVAFLSFIRKDSEETHLSGGYLDEL 118  
|||::||||| | : || :: | : | | | : | :  
Yr 333 LNLTSEKSQSKEVTETMKRKFCAQPMEELIKTSYG-VREKYLSECFSGTILSL 386  
  
RESULT 12 STANDARD; PRT; 175 AA.  
xx xx xxxxxx  
DE Sequence 11, Application US/08408095  
cc cc  
cc Sequence 11, Application US/08408095  
cc Patent No.5858678  
cc GENERAL INFORMATION:  
cc APPLICANT: Chinnadural, Govindaswamy  
cc TITLE OF INVENTION: APOPTOPOSIS-REGULATING PROTEINS  
cc NUMBER OF SEQUENCES: 35  
cc CORRESPONDENCE ADDRESS:  
cc ADDRESSEE: SUGHRUE, MILO, ZINN, MACPHEE & SEAS  
cc STREET: 2100 Pennsylvania Avenue, N.W.  
cc CITY: Washington  
cc STATE: D.C.  
cc COUNTRY: USA  
cc ZIP: 20037  
cc  
cc COMPUTER READABLE FORM:  
cc MEDIUM TYPE: Floppy disk  
cc OPERATING SYSTEM: IBM PC compatible  
cc SOFTWARE: PatentIn Release #1.0, Version #1.30  
cc CURRENT APPLICATION DATA:  
cc APPLICATION NUMBER: US/08/408,095  
cc FILING DATE: 21-MAR-1995  
cc CLASSIFICATION: 435  
cc ATTORNEY/AGENT INFORMATION:  
cc NAME: Mack, Susan J.  
cc REGISTRATION NUMBER: 30,951  
cc TELECOMMUNICATION INFORMATION:  
cc TELEPHONE: (202)293-7060  
cc FAX: (202)293-7860  
cc INFORMATION FOR SEQ ID NO: 11:  
cc SEQUENCE CHARACTERISTICS:  
cc LENGTH: 175 amino acids
```

```
CC      TYPE: amino acid
CC      STRANDEDNESS:
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
SQ      SEQUENCE 175 AA; 20396 MW; 142906 CN;

Query Match          2.8%; Score 93; DB 2; Length 175;
Best Local Similarity 29.1%; Pred. NO. 3.02e+01;
Matches 16; Conservative 13; Mismatches 24; Indels 2; Gaps 2

Db    65 INLHGQALFQEKVKTLD-FSTPGRAAAAVAFLEFPDKMSEETHLSGGYLLDPL 118
      |||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy    333 INTLSEKVSQEKVTMMKKFCQAQPMEIKTYSAG-VREKYLTSEYCFSGTYILSL 386

RESULT 13
ID     US-08-408-095-10        STANDARD; PRT; 175 AA.
XX     xxxxxx
DT
DE
XX
XX
Sequence 10, Application US/08408095
CC     Patent No. 5858678
CC     GENERAL INFORMATION:
CC     APPLICANT: Chindadural, Govindaswamy
CC     TITLE OF INVENTION: APOPTOSIS-REGULATING PROTEINS
CC     NUMBER OF SEQUENCES: 35
CC     CORRESPONDENCE ADDRESS:
CC     ADDRESSEE: SUGHRUE, MIOM, ZINN, MACPEAK & SEAS
CC     STREET: 2100 Pennsylvania Avenue, N.W.
CC     CITY: Washington
CC     STATE: D.C.
CC     COUNTRY: USA
CC     ZIP: 20037
CC     COMPUTER READABLE FORM:
CC     MEDIUM TYPE: Floppy disk
CC     COMPUTER: IBM PC compatible
CC     OPERATING SYSTEM: PC-DOS/MS-DOS
CC     SOFTWARE: PatentIn Release #1.0, Version #1.30
CC     CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER: US/08/408,095
CC     FILING DATE: 21-MAR-1995
CC     CLASSIFICATION: 435
CC     ATTORNEY/AGENT INFORMATION:
CC     NAME: Mack, Susan J.
CC     REGISTRATION NUMBER: 30,951
CC     TELECOMMUNICATION INFORMATION:
CC     TELEPHONE: (202)293-7060
CC     TELEFAX: (202)293-7860
CC     INFORMATION FOR SEQ ID NO: 10:
CC     SEQUENCE CHARACTERISTICS:
CC     LENGTH: 175 amino acids
CC     TYPE: amino acid
CC     STRANDEDNESS:
CC     TOPOLOGY: linear
CC     MOLECULE TYPE: protein
SQ     SEQUENCE 175 AA; 20369 MW; 143135 CN;

Query Match          2.8%; Score 93; DB 2; Length 175;
Best Local Similarity 29.1%; Pred. NO. 3.02e+01;
Matches 16; Conservative 13; Mismatches 24; Indels 2; Gaps 2;

Db    65 INLHGQALFQEKVKTLD-FSTPGRAAAAVAFLEFPDKMSEETHLSGGYLLDPL 118
      |||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy    333 INTLSEKVSQEKVTMMKKFCQAQPMEIKTYSAG-VREKYLTSEYCFSGTYILSL 386

RESULT 14
ID     US-08-408-095-9         STANDARD; PRT; 175 AA.
XX
```

```
AC      xxxxxx  
XX  
DT  
XX  
DE      Sequence 9, Application US/08408095  
XX  
CC      Sequence 9, Application US/08408095  
CC      Patent No. 585678  
CC      GENERAL INFORMATION:  
CC          APPLICANT : Chinadural, Govindasamy  
CC              TITLE OF INVENTION : APOPTOSIS-REGULATING PROTEINS  
CC          NUMBER OF SEQUENCES : 35  
CC      CORRESPONDENCE ADDRESS:  
CC          ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
CC              STREET: 2100 Pennsylvania Avenue, N.W.  
CC              CITY: Washington  
CC              STATE: D.C.  
CC              COUNTRY: USA  
CC              ZIP: 20037  
CC      COMPUTER READABLE FORM:  
CC          MEDIUM TYPE: Floppy disk  
CC      COMPUTER: IBM PC compatible  
CC      OPERATING SYSTEM: PC-DOS/MS-DOS  
CC      SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC      CURRENT APPLICATION DATA:  
CC          APPLICATION NUMBER: US/08/408,095  
CC          FILING DATE: 21-MAR-1995  
CC          CLASSIFICATION: 435  
CC          ATTORNEY/AGENT INFORMATION:  
CC              NAME: Mack, Susan J.  
CC              REGISTRATION NUMBER: 30,951  
CC          TELECOMMUNICATION INFORMATION:  
CC              TELEPHONE: (202)293-7060  
CC              TELEFAX: (202)293-7860  
CC          INFORMATION FOR SEQ ID NO: 9:  
CC      SEQUENCE CHARACTERISTICS:  
CC          LENGTH: 175 amino acids  
CC          TYPE: amino acid  
CC          STRANDEDNESS:  
CC          TOPOLOGY: linear  
CC      MOLECULE TYPE: protein  
CC      SEQUENCE 175 AA; 20402 MW; 143285 CN;  
  
Query Match           2.8%; Score 93; DB 2; Length 175;  
Best Local Similarity 29.1%; Pred. No. 3.02e+01;  
Matches   16; Conservative    13; Mismatches 24; Indels    2; Gaps    2;  
  
Db       65 LNLGQALFGQEVIRITLD-FSTPRGAANAVALFSIKRKMSSETHLSGYLIDFL 118  
        |||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
QY     333 LNLTSKVSQEKVETEMKRKFCAQPWEIIRTYSAG-VKEKYISEYCFCGTYYLSLT 386  
  
RESULT  15  
ID      US-08-530-792D-23 STANDARD: PRT: 1025 AA..  
XX      xxxxxx  
XX  
DT  
XX  
DE      Sequence 23, Application US/08530792D  
XX  
CC      Sequence 23, Application US/08530792D  
CC      Patent No. 5972680  
CC      GENERAL INFORMATION:  
CC          APPLICANT: Knowles, W.J.; Guralski, D.; Haigh, W.; Letsinger, J.T., et al.  
CC              TITLE OF INVENTION: Glucose Transporter Vesicle Aminopeptidase  
CC          NUMBER OF SEQUENCES: 23  
CC      CORRESPONDENCE ADDRESS:  
CC          ADDRESSEE: Bayer Corporation  
CC              STREET: 400 Morgan Lane  
CC              CITY: West Haven  
CC              STATE: Connecticut
```

```

CC      COUNTRY :   U.S.A.
CC      ZIP:      06516
CC
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: 3.5" diskette, 1.44 MB Storage
CC      COMPUTER : Dell Windows 95 PC
CC      OPERATING SYSTEM: Windows 95
CC      SOFTWARE: Wordperfect for Windows 6.1
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/530,792D
CC      FILING DATE: 09/19/95
CC      CLASSIFICATION: 435
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: 08/309,232
CC      FILING DATE: 09/20/94
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Brewer, Alice A.
CC      REGISTRATION NUMBER: 32888
CC      REFERENCE/DOCKET NUMBER: MW 323P1
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (203) 812-2705
CC      TELEFAX: (203) 812-5492
CC      INFORMATION FOR SEQ ID NO: 23:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 1025 amino acids
CC      TYPE: amino acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein;
CC      ORIGINAL SOURCE:
CC      ORGANISM: Rattus norvegicus
CC      STRAIN: Sprague-Dawley
CC      DEVELOPMENTAL STAGE: adult
CC      TISSUE TYPE: skeletal muscle
CC      IMMEDIATE SOURCE:
CC      LIBRARY: Clontech rat skeletal muscle cDNA library in lambda
CC      CLONE: 12-1 (from lambda gtl1 library, PCR product clones 5,
CC      CLONE: 334, and KC44.
CC      FEATURE:
CC      NAME/KEY: complete amino acid sequence for GTVap, long version
CC      IDENTIFICATION METHOD: translation from cDNA
CC      SEQUENCE 1025 AA; 117200 MW; 5643821 CN;
SQ
Query Match 2.8%; Score 92; DB 2; Length 1025;
Best Local Similarity 25.8%; P22; Mismatches 36; Indels 8; Gaps 8;
Matches 23; Conservative
Db 485 INEGCATPMEFFSEVKEIFKEINSEYEDFLDARFKTKMRKDSLNSHPISISSVOSS-EQIEEM 543
Oy 316 LQGDGCAISATYFPWK-FLNLTLS-EKVSQEKVTEMKK-F-CAP-WEIKTYSAGYKER 370
Db 544 FDSLSEYFKASLL-LMKSY-LSEDEYFOH 570
Oy 371 YLSEYCFSGFYILLSLQGYHFTADSWEH 399

```





QY	RESULT	ENTRY	2
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	#type complete
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	hypothetical protein R07E4.4 - Caenorhabditis elegans
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	#formal_name Caenorhabditis elegans
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	#text_change 20-Sep-1999 #sequence_revision 20-Sep-1999
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	#accession 116696
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	#authors Miller, N.
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	#submission submitted to the EMBL Data Library, October 1995
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	#description The sequence of C. elegans cosmid R07E4.
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	#status preliminary; translated from GB/EMBL/DBJ
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	#molecule_type DNA
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	#residues 1-557 #label MITL
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	#cross_references EMBL:039652; NID:g1049390; PID:g1049394;
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	PIDN:AAA80403.1; CESP:R07E4.4
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	#experimental_source strain Bristol N2
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	CESP:R07E4.4
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	39/1; 67/1; 145/2; 192/1; 244/3; 317/3; 451/3; 488/2
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	#length 557 #molecular_weight 63056 #checksum 4267
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	Query Match 11.2% Score 366; DB 2; Length 557;
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	Best Local Similarity 25.6% Pred. No. 1,19e-47;
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	Matches 111; Conservative 109; Mismatches 160; Indels 34; Gaps 27
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	44 YGVLDAGSTGRFLVYWMISTSDSELIQIPVIYDKNPVKKIKISPGLSIFPTKPAQAAE 103
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	12 YGVLDAGSTGRFLVYWMISTSDSELIQIPVIYDKNPVKKIKISPGLSIFPTKPAQAAE 103
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	104 YLRPLMAELAKRIHEPEKRPYVFIFFATAGKRLIPDEYVLIGQKAVLYAKNLRNKLPKITS 163
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	67 YLTDCMERAREVIEPRSOHQETPPVYLGATAGRLLMESEELADR-VLDYVERSLSNY-P 123
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	164 MOVLKEHRIIEGKKEGYSIMVAVYALGKRNKATLDPETSPAHNAOKTVGMIDMGA 223
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	124 PD-FQG-ARIIITGQEGAVGMITTYLTKGKSQKTR-FSTVPEPTNNQEFFGALDDGA 180
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	224 SAQILFELPDITDSESSIVENINIGCREDDSLFYKFLVTFILGYGVNEGIRKEHMLLS 283
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	181 STQYTF-VPOQNTIES-P-DN-ALQFLYGR-DYNYVTHSFLCYGKDQALM---QRLAK 231
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	284 KIKDQNGTVIODDCMPLNLAKTVTLENG-ENF-VRRGTGNNTGSNEYKLLN-PSSSE 340
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	232 DIQVASNEITLRDPCFHPKRYKVNVSDLYCTPKTRKREMTLPFOOFETIGTIGNYQCHQS 291
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	341 VCKA-EAAKCYFGAVAPASIPLSINEM-YG-FSEKWTSTHYVLGLGQ-YAENIAKTQ 396
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	292 ILELENTSYCYSOCAFNGIFLPLLOGGFGAFSAFYVMK-FLNLTSEKYSQKEVTMMK 350
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	397 QYKSRMTIOAESKOLYPRADERLRTQOFSKAMITSVLHDGS-VDKTHNKFQSVST 455
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	351 KFCQAPW---E-EIKTS-YAGVKEKYLSEYFSGTYIILSLLOGYHFTADSWEHIFHTGK 405
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	456 IAGQEVQWALGAMI 469
301	CPYSCAANGIFLPPLOQDFGAFSAFYVVMFLNLTSSKVOEKTEEMKKFFCAQNPWEI	116696	406 IQGSDAGWTGLGYML 419

ENTRY	550463	#type complete
TITLE	hypothetical protein YER005w - yeast (Saccharomyces cerevisiae)	
ORGANISM	#format_name Saccharomyces cerevisiae	
DATE	28-May-1993 #sequence_revision 24-Feb-1995 #text_change 21-Nov-1997	
ACCESSIONS	S50463	
REFERENCE	S50428	
#authors	Dietrich, F.S.	
#submission	Submitted to the EMBL Data Library, December 1994	
#description	The sequence of S. cerevisiae cosmid 9537, 9581, 9495, 9867 and lambda clone 5898.	
#accession	S50463	
#molecule_type	DNA	
#residues	1-630 #label DIE	
#cross-references	EMBL:U018778; NID:g603592; PID:g603597; MIPS:YER005w	
GENETICS		
#map_position	5R	
SUMMARY	#length 630 #molecular_weight 71851 #checksum 7317	
Query Match	11.0%; Score 360; DB 2; Length 630;	
Best Local Similarity	28.4%; Pred. No. 1,89e-46;	
Matches	106; Conservative 93; Mismatches 137; Indels 37; Gaps 28;	
Db	52	IHKQKDWTFKINPGIASSPEKKRDQAYKSHIKPLDPFAKNIIPESHSWSSCFVFIQATGMR 111
QY	41	VHVEECRVR-GPGRISKVKQVNEI-GIYLDCKERAREVTPRSQHOETPVYLGATGMR 98
Db	112	LLPDIQSSILIDGICQGLKHP-AFIVEDESAQIQVIDGETEGLYGMLNYLVGHFND- 169
QY	99	LLRMESR-LADRLVDYERSLSNYP-DRGA-RITGQEGAYGITTINYLGRSQK 155
Db	170	--YN-P-EVSDHFTTGFMDSGASTQIAFAPHDSEIARHRDIAITIFLRVNSDQK 223
QY	156	TRMFESIVPEYENNOETGALDLGGASTQVFEVPONQ-RESP-DN-ALQF-R-LyGK-D- 208
Db	224	WDVSVSTWLGAGNAQARRRYLAQILNLPENT-ND-YENDEFSIR-NLNDPCPRGSSTD 280
QY	209	YNYTHSLTLCGKQALQMLAKIOVASNIELRDPCHFGYKRVVNSDLYRTP-CTKR 267
Db	281	PEFKDTI-PIHAGSGVNEOCTKSIYPLLLKNMPCDEPCLFNGVHAPRIDFANDKFIGT 338
QY	268	PEMLPPOQFPIQIGTIGNTQCHQSILTELPNTSY-CPISQCAFNKIFLPLQ-G-D-PGAR 323
Db	339	SEWYTANDYFKLGE-VNEDKFSKSLREFCNSNMTQILANSKGYVNSIPENFLDACP 397
QY	324	SAFYFVKM-FLNLTSEKVSQEKVTEMMKKRCAQWEEI-----KTSYAGVKEKYLSEYCF 377
Db	398	KGNWVNLINHEGF 410
QY	378	SGTYILSLILQY 390
RESULT	4	
ENTRY	JC4616	#type complete
TITLE	aypase (EC 3.6.1.5) precursor - potato	
ALTERNATE_NAMES	adenylpyrophosphatase; ATP-diphosphohydrolase	
ORGANISM	#format_name Solanum tuberosum #common_name potato	
DATE	10-May-1996 #sequence_revision 19-Jul-1996 #text_change 28-May-1999	
ACCESSIONS	JC4616; PC4147	
REFERENCE	JC4616	
#authors	Handa, M.; Guidotti, G.	
#journal	Biochem. Biophys. Res. Commun. (1996) 218:916-923	
#title	Purification and cloning of a soluble ATP-diphosphohydrolase (Apyrase) from potato tubers (Solanum tuberosum).	
#cross-references	MUTID:96158985	
#accession	JC4616	
#molecule_type	mRNA	
#residues	1-454 #label HAN	
#cross-references	GB:U58597; NID:g1381632; PIDN:AA02720.1; PID:g1381633	
#accession	PC4147	



[illegible]

KEYWORDS glycoprotein; Golgi apparatus; hydrolase; transmembrane protein

FEATURE 10-24 #domain transmembrane #status predicted #label TM\

41,280,335 #binding\_site carbohydrate (Asn) (covalent) #status predicted

SUMMARY #length 518 #molecular-weight 56821 #checksum 4379

Query Match 8.1%; Score 265; DB 2; Length 518;

Best Local Similarity 27.8%; Pred. No. 6,47e-28; Mismatches 180; Indels 45; Gaps 37;

Matches 122; Conservative 92; Mismatches 180; Indels 45; Gaps 37;

Db 88 SEHKYIMIDAGSTGRVHYKFDCTSPPT-LDEKPD-MLE-PGLSSFDSDYGAAN 144

QY 7 PENVKKGIYLDAGSSHTSLYIKWPAKENDTGVAHQVEECRVKGCISKF-YQKVAEIG 65

Db 145 S-LDPLIKAMNVPIKANSCTPVAVKATAGLGLDASSKILSANRDLKEDYR-FPV 202

QY 66 IYLDCEMERAREVIPSQOETFPVYLGATAGMRLT-RMESEELADRLVDVERSLSNYPF 124

Db 203 VEGDGSINGSGDEGEVAFMTTNYLGNIGANGP--KL-P-TAA--VE---DLGGGSTQ 252

QY 125 -DPGARITTGDEBAGVMTTNYLGNIGANGP--KL-P-TAA--VE---DLGGGSTQ 193

Db 253 IVEEPTPIERKAVDEHREKFDLKGDENTLYXQSHLGYLKEG-RNKV-NSVLVE-NA- 308

QY 184 VFEVPOQNTIES-PDNALQFRL-YGKD-YNVYHSLFCGKQDALMQLAKDIQVANSNEI 240

Db 309 LKDGLIKGDMTKTHOLSPCLPPKVNATNEKYLTSKEYTITDFIGPDESSGACRELT 368

QY 241 LBDP-CFHGKRVKVVNSDLYKTP-CTKRFE-MTLRFPQ-FEIOGIGNYQC-HQS-IL- 293

Db 369 DELINKDACQSPGSCFNGHPSLVRTFKESNDIYFSEYFRTPLGMLSFITNELN 428

QY 294 -ELENT-STCPYSCAFNGITFLPLQGD--GA-FSAF-YFVKFLNLT-SEKVSQEKVT 346

Db 439 DLARIYCKGEETWNSVSGISIDLESDFHCLDLSEFVSLHTGYDIPLO-RE-LRT 486

QY 347 EMKKKFC-AQ-PWEIKTYSAGVKEKYLSE--YCSFGTYLSLLQGYHTADSWEHIF 402

Db 487 GKIKKEIGKCLGASLPL 505

QY 403 ICKIGSDAGMTLGYML 421

RESULT 7 #type complete

ENTRY T04439 #intron 1

TITLE hypothetical protein T18B16.150 - Arabidopsis thaliana

ORGANISM #formal\_name Arabidopsis thaliana #common\_name mouse-ear

DATE 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change

ACCESSIONS T04439

REFERENCE 215359

#authors Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.;

#submission submitted to the Protein Sequence Database, April 1998

#accession T04439

#molecule\_type DNA

#residues 1-1052 #label BEV

#cross-references EMBL:AL021687

#experimental\_source cultivar Columbia; BAC clone T18B16

GENETICS #map\_position 4

#intron 79/3; 118/3; 217/1; 295/3; 396/3; 531/2; 815/3

SUMMARY #note T18B16.150

#length 1052 #molecular-weight 116388 #checksum 4955

Query Match 8.1%; Score 266; DB 2; Length 1052;

Best Local Similarity 27.5%; Pred. No. 4.18e-28; Mismatches 104; Conservative 103; Mismatches 125; Indels 46; Gaps 33;

QY 52 PGISKEVQKVNIEGIVYLDCEMERAREVIPSQOETFPVYLGATAGMRLT-RMESEELADRV 111

Db 635 LGVNMILAKSPPTCRREWKITSGTEEAFFGWTALNT-----QTSN-LGLP---KK- 683

QY 112 LDVERSLSNYPD-FQG-ARITGOEGAYGMITTNYLGNIGANGP--KL-P-TAA--VE---DLGGGSTQ 169

Db 684 ATFGALDGLSSLOVTFEERHNETNLTNRIGSVNHLISVSLAGYGLNADPDRSVH 743

QY 170 ETFGALDGLGASQVTFVFNQNIIESPDNALPRLXKRYNTYS-F-L-CYKQDA- 224

Db 744 LKRLPNVKSDDLIEKLEKHPCLNSGVN-GQYCSQC-ASSVQGGKSGVSIKLVG 801

QY 225 LMOKLAK-DI-QVASNEI-LRDCCFHFGYKRVVNSDLYKTPCTKRFEMTLPLFOQFEIOG 281

Db 802 APWAGEC--SALAK-NAP-C--AL-P-DG-YPRP-HGQYVANSFGFYVRFNLSAF-AS 850

QY 282 IGNYQOCHOSTIELEFNTSYCPYSCAFNGITFLPLQGDGAFSAFVFWKFLNLTSEKVS 341

Db 851 LDVLEKREPCDKAMQVARTS-VSP-QPFIQCFRAPHYVSLRREGT-YITDK--QI- 904

QY 342 QEKVTEMKKFKCAQPEBEIKTYSAGVKEKYLSEYCSGYTILSLQGYHTADSWEHIF 401

Db 905 IIGS--GSIT-WTLGYAL 919

QY 402 FICKIGSDAGMTLGYML 419

RESULT 8 #type complete

ENTRY T05213 #intron 1

TITLE hypothetical protein F1715.40 - Arabidopsis thaliana

ORGANISM #formal\_name Arabidopsis thaliana #common\_name mouse-ear

DATE 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change

ACCESSIONS T05213

REFERENCE 215404

#authors Bevan, M.; Vitale, D.; Laguori, R.; Argiridou, A.; De Simone,

#submission submitted to the Protein Sequence Database, July 1998

#accession T05213

#molecule\_type DNA

#residues 1-371 #label BEV

#cross-references EMBL:AL031032

#experimental\_source cultivar Columbia; BAC clone F1715

GENETICS #map\_position 4

#intron 14/1; 43/3; 117/1; 171/3

SUMMARY #note F1715.40

#length 371 #molecular-weight 42304 #checksum 9483

Query Match 4.0%; Score 131; DB 2; Length 371;

Best Local Similarity 23.8%; Pred. No. 1.16e-04; Mismatches 19; Conservative 29; Mismatches 26; Indels 6; Gaps 6;

Db 197 HOBOSYRKARKAVRIRAVSEGOPIPNATISIOQKRLG-PFGCEVEKNILG-NQ-AYE 253

QY 84 HOETPYTGATAGMRLNIESE-E-LADRYLDIVERSLSNPFDPGARRITGOEGAY- 140

Db 254 NMFOTRYTFPANEKWS 273

QY 141 GMTTNYLGNIGANGP--KL-P-TAA--VE---DLGGGSTQ 169

RESULT 9 #type complete

ENTRY A55421 #intron 1

TITLE nucleoside-triphosphatase (EC 3.6.1.15) 1 precursor -

ORGANISM #formal\_name Toxoplasma gondii

DATE 10-Feb-1995 #sequence\_revision 10-Feb-1995 #text\_change

ACCESSIONS A55421

REFERENCE A55421

```

#authors      Bermudes, D.; Peck, K.R.; Affili Affili, M.; Beckers, C.J.M.;
#journal      J. Biol. Chem. (1994) 269:29252-29260
#title        Tandemly repeated genes encode nucleoside triphosphate
#             hydrolyase isoforms secreted into the parasitophorous
#             vacuole of Toxoplasma gondii.
#cross-references MUID:95050750
#accession    A55421
##status      preliminary
##molecule_type DNA
##residues    1-628 ##label BER
##cross-references GB:U14322; NID:g562078; PID:g562079
KEYWORDS      hydrolase
SUMMARY       #length 628 #molecular-weight 69585 #checksum 5865

Query Match      3.9%; Score 127; DB 2; Length 628;
Best Local Similarity 32.8%; Pred. No. 4.53e-04;
Matches 22; Conservative 21; Mismatches 22; Indels 2; Gaps 2;

Db 228 TRPTGAEGLFAFTTINLSTRIGEDPARCMIDEXGVHCHNDLAGVYEVGASAGQIVF 287
Qy 129 ARITGQEGAGYGTITINLLGKFSQKTRWFSTVPEYETNN-OETF-GALDLAGASTQYTF 186

Db 288 PLOEGTV 294
Qy 187 VPQNOTI 193

RESULT 10
ENTRY      T04954 #type complete
TITLE      hypothetical protein F7J7.170 - Arabidopsis thaliana
ORGANISM   #formal_name Arabidopsis thaliana #common_name mouse-ear
            cress
DATE        23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change
            11-Jun-1999
ACCESSIONS T04954
REFERENCE   215391
#authors    Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.;
            Mewes, H.W.; Mayer, K.F.X.; Schueler, C.
#submission submitted to the Protein Sequence Database, July 1998
#accession  T04954
##molecule_type DNA
##residues  1-597 ##label BEV
##cross-references EMBL:AL021960
##experimental_source cultivar Columbia; BAC clone F7J7

GENETICS
#map_position 4
#introns      266/1; 311/1; 349/3; 454/2; 503/3
#note         F7J7.170
SUMMARY       #length 597 #molecular-weight 66904 #checksum 5095

Query Match      3.8%; Score 125; DB 2; Length 597;
Best Local Similarity 26.8%; Pred. No. 8.86e-04;
Matches 26; Conservative 28; Mismatches 38; Indels 5; Gaps 5;

Db 412 RDLKSSNVLDEQMLPKISDFGMAR-QEDEFDTQAVTRRRVGTGYMAPEYMHGRFSYK 470
Qy 98 RLLRMESEELADRLV-DVVERSLSNYPFDFOGARITIGQESAGYGMITINYL-KGFSQK 155

Db 471 TDVYSFGVLYLEITGKRNSGGLGEGTDLPPEAMON 507
Qy 156 TR-W-FSTVPEYETNNQETFGALDLGAGASTQYTFPON 190

RESULT 11
ENTRY      JN0620 #type complete
TITLE      UDP-glucuronosyltransferase (EC 2.4.1.-) 2B-10 precursor -
            human
ORGANISM   #formal_name Homo sapiens #common_name man
            human
DATE        31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
            28-May-1999
ACCESSIONS JN0620
REFERENCE   JN0619

```

```

#authors      Jin, C.J.; Miners, J.O.; Lillywhite, K.J.; Mackenzie, P.I.
#journal      Biochem. Biophys. Res. Commun. (1993) 194:496-503
#title        cDNA cloning and expression of two new members of the human
#             liver UDP-glucuronosyltransferase 2B subfamily.
#cross-references MUID:93326164
#accession    JN0620
##molecule_type mRNA
##residues    1-528 ##label JIN
##cross-references GB:X63359; NID:g516149; PIDN:CAAA4961.1; PID:g516150
CLASSIFICATION #experimental_source liver
                #superfamily glucuronosyltransferase
KEYWORDS      glycoprotein; glycosyltransferase; hexosyltransferase;
                transmembrane protein

FEATURE
1-23          #domain signal sequence #status predicted #label SIG\
24-528        #product UDP-glucuronosyltransferase 2B-10 #status
                predicted #label MET\
491-508        #domain transmembrane #status predicted #label TM\
66,314,481     #binding_site carbohydrate (Asn) (covalent) #status
                predicted

SUMMARY       #length 528 #molecular-weight 60773 #checksum 4903

Query Match      3.6%; Score 117; DB 2; Length 528;
Best Local Similarity 24.7%; Pred. No. 1.22e-02;
Matches 21; Conservative 24; Mismatches 36; Indels 4; Gaps 4;

Db 133 KRLMKLOESRPDIYPADAYLPCGELLAELEFNIPV-YSH-SFSPQSYERRSGCF-IFP 189
Qy 266 KRFEMLLPQGFQEIQIGNGYQCHQSIIELEFMTSCYPCSQAFNGIF-LPPIQGFAGPS 324

Db 190 PSYVYVWMSKLSQDQTFMERYKNML 214
Qy 325 AFYFWKFLNLTSEKVSQEKVTEMM 349

RESULT 12
ENTRY      G69998 #type complete
TITLE      thiorodoxin H1 homolog yfcp - Bacillus subtilis
ORGANISM   #formal_name Bacillus subtilis
DATE        05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
            11-Jun-1999
ACCESSIONS G69998
REFERENCE   A69580
#authors    Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
            Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.;
            Bojotin, A.; Borcherdt, S.; Boriss, R.; Boursier, L.; Brans,
            A.; Braun, M.; Brigelli, S.C.; Bron, S.; Brouillier, S.;
            Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.;
            Choi, S.K.; Codani, J.J.; Conerton, I.F.; Cummings, N.J.;
            Daniel, R.A.; Denizot, F.; Devine, K.M.; Diesterhoelt, A.;
            Ehrlich, S.D.; Emerson, P.T.; Entlian, K.D.; Erlington, J.;
            Fabret, C.; Ferrari, E.; Foulger, D.; Frit, C.; Fujita,
            M.; Fujita, Y.; Fuma, S.; Galliz, A.; Galleron, N.; Ghim,
            S.Y.; Glaser, P.; Goffeau, A.; Gollightly, E.J.; Grandi, G.;
            Guisepi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
            C.R.; Henaut, A.; Hlibert, H.; Holsappel, S.; Hosono, S.;
            Hullo, M.F.; Itaya, M.; Jones, J.; Joris, B.; Karamata, D.;
            Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi,
            Y.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.;
            Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.;
            Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
            Mameel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno,
            M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly,
            M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro,
            V.; Pohl, T.M.; Portetelle, D.; Porrolik, S.; Prescott,
            A.M.; Presecan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.;
            Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocho, E.;
            Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, E.;
            Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.;
            Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo,
            B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.;
            Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.;
            Tepstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.;

```

		Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.; Wandut, R.; Wedler, E.; Wedler, H.; Weitzengger, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A.
#journal		Nature (1997) 390:249-256
#title		The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
#cross-references		NCBI:98044033
#accession		G69998
#status		preliminary; nucleic acid sequence not shown; translation not shown
##molecule_type	DNA	
##residues	1-107	##label KUN
##cross-references	GB:Z99119; GB:AL009126; NID:g2635411; PIDN:CAB14962.; PID:e1185857; PID:g2635468	
##experimental_source	strain 168	
GENETICS		
#gene	ycpP	
CLASSIFICATION	#superfamily thioedoxin; thioedoxin homology	
SUMMARY	#length 107 #molecular_weight 12747 #checksum 1509	
Query Match	3.5%; Score 116; DB 2; Length 107;	
Best Local Similarity	34.8%; Pred. No. 1.68e-02;	
Matches	16; Conservative 13; Mismatches 12; Indels 5; Gaps 4;	
Db	22 MFSADWCP--DCRFVEFLPELEANEPEPF--YYVDKRTIDTCAG 64 :::     :   :	
Oy	295 LNTSYCPYSCAFNGICIFLPPLGDFGCAFAFYFV-M-KELNLTSR 338	
RESULT	13	
ENTRY	T12886	#type complete
TITLE	conserved hypothetical protein yora - Bacillus subtilis phage SphC2	
ORGANISM	#formal name Bacillus subtilis phage SphC2	
DATE	13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 20-Aug-1999	
ACCESSIONS	T12886; B69922	
REFERENCE	Lazarovic, V.; Duesterhoef, A.; Soldo, B.; Hilbert, H.; Manel, C.; Karamata, D.	
#authors	submitted to the EMBL Data Library, August 1997	
#submission	The complete nucleotide sequence of the Bacillus subtilis SphC2 prophage.	
#description	T12886	
#accession	preliminary; translated from GB/EMBL/DDBJ	
##status	DNA	
##molecule_type	DNA	
##residues	1-586 ##label LAZ	
##cross-references	EMBL:AF020713; NID:g3025478; PID:g3025600; PIDN:AC13095.1	
REFERENCE	A69580	
#authors	Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.; Bocting, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans, A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani, J.J.; Conerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoef, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Erington, J.; Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim, S.Y.; Glaser, P.; Goftau, A.; Golightly, E.J.; Grandi, G.; Guiseppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hulio, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.; Kashnira, Y.; Klair-Blanchard, M.; Klein, C.; Kobayashi, Y.; Koeltter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidas, A.; Iadinoie, A.; Liu, H.; Masuda, S.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maneel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno, M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogiwara, A.; Onodera, B.; Park, S.H.; Parro,	

	#journal	The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.							
	#title	Nature (1997) 390:249-256							
	#cross-references MIMD:	B69922							
	#accession	B69922	nucleic acid sequence not shown; translation not shown						
	#status	##molecule_type DNA							
	#residues	1-586 ##label KUN							
	#cross-references GB:	Z99114; CB:AL009126; MID:g2634230; PID:e185517; PID:g2634438							
	GENETICS	##experimental_source strain 168							
SUMMARY	#gene yorA	length 586 #molecular-weight 64972 #checksum 7245							
	Query Match	3.3%; Score 109; DB 2; Length 586;							
	Best Local Similarity	31.1%; Pred. No. 1,50e-01;							
	Matches 14; Conservative	13; Mismatches 17; Indels 1; Gaps 1;							
Db	506 SEHWARGNEIMFNDRSDGAGIYNGKRYKTRINDVGIGAVIL	550							
Oy	69 TDCMERAREVI-PRSOHOETPVYLGAATGAKMLRMESEELADRYL	112							
RESULT 14	T05212	#type complete							
ENTRY	hypothetical protein FL1715_30 - Arabidopsis thaliana								
TITLE	#formal_name Arabidopsis thaliana tcommon_name mouse ear cross								
ORGANISM	DNA								
DATE	23-Apr-1999	#sequence_revision 23-Apr-1999	#text_change						
ACCESSIONS	T05212								
REFERENCE	215404								
#authors	Bevan, M.; Vitale, D.; Liguori, R.; Argirou, A.; De Simone, V.; Hohenseil, J.; Meves, H.W.; Mayer, K.F.X.; Scheller, C								
#submission accession	T05212								
##molecule_type DNA									
##residues	1-669 ##label BEV								
#cross-references EMBL:	AL011032								
##experimental_source cultivar Columbia; BAC clone FL1715									
GENETICS									
#map_position 4	60/3; 125/3; 199/1; 253/3; 549/3								
#introns	FL1715_30								
SUMMARY	length 669 #molecular_weight 75206 #checksum 5116								
	Query Match	3.3%; Score 108; DB 2; Length 669;							
	Best Local Similarity	23.2%; Pred. No. 2.03e-01;							
	Matches 16; Conservative	21; Mismatches 28; Indels 4; Gaps 4;							
Db	250 GTVRIRVVNNKGITIPNATISIEBKRG-YPCGCAYENNLLS-NQ-AYNMFOTGFVTYT	346							
Oy	93 ATAGARILLRMESSELADVDLVERSLSNPDPQGARIITGGEBAG-GMITINPLDGK	151							
Db	347 FGNEKKWYS	355							
Oy	152 FSGKRWFWS	160							

```

RESULT 15
ENTRY S63501 #type fragments
TITLE apyrase (EC 3.6.1.5) - human (fragments)
ALTERNATE_NAMES ATP diphosphohydrolase
ORGANISM #formal_name Homo sapiens #common_name man
DATE 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change
13-Mar-1997
ACCESSIONS S63501
REFERENCE S63501
#authors Christoforidis, S.; Papamarcaki, T.; Galaris, D.; Kellner,
R.; Tsolas, O.
#journal Eur. J. Biochem. (1995) 234:66-74
#title Purification and properties of human placental ATP
diphosphohydrolase.
#cross-references MUID:96096723
#accession S63501
#status preliminary
#molecule_type protein
#residues 1-3;4-8;9-13;14-20;21-30;31-44 #label CHR
KEYWORDS hydrolase
SUMMARY #length 44 #checksum 6585

Query Match 3.2%; Score 104; DB 2; Length 44;
Best Local Similarity 51.3%; Pred. No. 6,75e-01;
Matches 20; Conservative 6; Mismatches 10; Indels 3; Gaps 2;

Db 9 LLSKFG-KD--YNYKKVYVSDLGALDLGASTQVTF 44
|||||:|::: : ||||| ||||| |||||
QY 148 LIGKFSQKRWFSIVPEYETNNQETFGALDLGASTQVTF 186

```

Search completed: Fri May 5 09:23:35 2000  
 Job time : 307 secs.





Db 241 LRDPCFHPGKRVVNSDLYKTPCTKRFEMTLTPOQFEIOGIGNYOQCHOSILELFTNTSY 300  
| | | | |  
Qy 241 LRDPCFHPGKRVVNSDLYKTPCTKRFEMTLTPOQFEIOGIGNYOQCHOSILELFTNTSY 300  
| | | | |  
Db 301 CPYSQCAFNGIFLPLDGDGAFSAFYVWKFLNLTSEKVSQEKVTEMMKKCAQPMWEI 360  
| | | | |  
Qy 301 CPYSQCAFNGIFLPLDGDGAFSAFYVWKFLNLTSEKVSQEKVTEMMKKCAQPMWEI 360  
| | | | |  
Db 361 KTSYAGVKKERYLSEYCFSGYIILSLLOGYHFTADSMENHIFIGKIQGSDAGWTGMYLN 420  
| | | | |  
Qy 361 KTSYAGVKKERYLSEYCFSGYIILSLLOGYHFTADSMENHIFIGKIQGSDAGWTGMYLN 420  
| | | | |  
Db 421 LTNMIPAEQPLSTPLSHST 439  
| | | | |  
Qy 421 LTNMIPAEQPLSTPLSHST 439  
| | | | |  
RESULT 2 STANDARD; PRT; 454 AA.  
ID PCT-US99-22955-6  
XX xxxxxx  
AC  
XX  
DT  
XX  
DE  
XX  
Sequence 6, Application PC/TUS9922955  
CC GENERAL INFORMATION:  
CC APPLICANT: Maliszewski, Charles R.  
CC APPLICANT: Gayle III, Richard B.  
CC APPLICANT: Price, Virginia L.  
CC APPLICANT: Gimpel, Steven D.  
CC APPLICANT: Immunex Corporation  
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment  
CC FILE REFERENCE: 2879-WO  
CC CURRENT APPLICATION NUMBER: PCT/US99/22955  
CC EARLIER FILING DATE: 1999-10-13  
CC EARLIER APPLICATION NUMBER: US 60/104,585  
CC EARLIER FILING DATE: 1998-10-16  
CC EARLIER APPLICATION NUMBER: US 60/107,466  
CC EARLIER FILING DATE: 1998-11-06  
CC EARLIER APPLICATION NUMBER: US 60/149,010  
CC EARLIER FILING DATE: 1999-08-13  
CC NUMBER OF SEQ ID NOS: 31  
CC SOFTWARE: Patentln Ver. 2.0  
CC SEQ ID NO 6  
CC LENGTH: 454  
CC TYPE: PRT  
CC ORGANISM: Artificial Sequence  
CC OTHER INFORMATION: construct of human CD39  
CC SEQUENCE 454 AA; 51543 MW; 1133800 CN;  
Query Match 100.0%; Score 3275; DB 1; Length 454;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 16 TONKALPENKYGIVLDAGSSHTSLYIKWPAKENDTGVAHVQVEBCRVKPGISKPVOK 75  
| | | | |  
Qy 1 TONKALPENKYGIVLDAGSSHTSLYIKWPAKENDTGVAHVQVEBCRVKPGISKPVOK 60  
| | | | |  
Db 76 VNEIGIYLTDCMERAREVIPRSOHOETPYVLTAGTMRLLRMESEELADRVLDYVERSLS 135  
| | | | |  
Qy 61 VNEIGIYLTDCMERAREVIPRSOHOETPYVLTAGTMRLLRMESEELADRVLDYVERSLS 120  
| | | | |  
Db 136 NYPFDFOGARIIITGOEGAGYGMTITNYLLGKFSQKTRWFSIVPYETNNQETFGALDLGA 195  
| | | | |  
Qy 121 NYPFDFOGARIIITGOEGAGYGMTITNYLLGKFSQKTRWFSIVPYETNNQETFGALDLGA 180  
| | | | |  
Db 196 STQVTFVPQNOTIESPDNALQFRLYGKDYNVYTHSFLCYGKDQALMOKLAKIOVASNEI 255  
| | | | |  
Qy 181 STQVTFVPQNOTIESPDNALQFRLYGKDYNVYTHSFLCYGKDQALMOKLAKIOVASNEI 240  
| | | | |  
Db 256 LBDPCFHPGKRVVNSDLYKTPCTKRFEMTLTPOQFEIOGIGNYOQCHOSILELFTNTSY 315  
| | | | |

Qy 241 LRDPCFHPGKRVVNSDLYKTPCTKRFEMTLTPOQFEIOGIGNYOQCHOSILELFTNTSY 300  
| | | | |  
Db 316 CPYSQCAFNGIFLPLDGDGAFSAFYVWKFLNLTSEKVSQEKVTEMMKKCAQPMWEI 375  
| | | | |  
Qy 301 CPYSQCAFNGIFLPLDGDGAFSAFYVWKFLNLTSEKVSQEKVTEMMKKCAQPMWEI 360  
| | | | |  
Db 376 KTSYAGVKKERYLSEYCFSGYIILSLLOGYHFTADSMENHIFIGKIQGSDAGWTGMYLN 435  
| | | | |  
Qy 361 KTSYAGVKKERYLSEYCFSGYIILSLLOGYHFTADSMENHIFIGKIQGSDAGWTGMYLN 420  
| | | | |  
Db 436 LTNMIPAEQPLSTPLSHST 454  
| | | | |  
Qy 421 LTNMIPAEQPLSTPLSHST 439  
| | | | |  
RESULT 3 STANDARD; PRT; 463 AA.  
ID PCT-US99-22955-30  
XX xxxxxx  
AC  
XX  
DT  
XX  
DE  
XX  
Sequence 30, Application PC/TUS9922955  
CC GENERAL INFORMATION:  
CC APPLICANT: Maliszewski, Charles R.  
CC APPLICANT: Gayle III, Richard B.  
CC APPLICANT: Price, Virginia L.  
CC APPLICANT: Gimpel, Steven D.  
CC APPLICANT: Immunex Corporation  
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment  
CC FILE REFERENCE: 2879-WO  
CC CURRENT APPLICATION NUMBER: PCT/US99/22955  
CC EARLIER FILING DATE: 1999-10-13  
CC EARLIER APPLICATION NUMBER: US 60/104,585  
CC EARLIER FILING DATE: 1998-10-16  
CC EARLIER APPLICATION NUMBER: US 60/107,466  
CC EARLIER FILING DATE: 1998-11-06  
CC EARLIER APPLICATION NUMBER: US 60/149,010  
CC EARLIER FILING DATE: 1999-08-13  
CC NUMBER OF SEQ ID NOS: 31  
CC SOFTWARE: Patentln Ver. 2.0  
CC SEQ ID NO 30  
CC LENGTH: 463  
CC TYPE: PRT  
CC ORGANISM: Artificial Sequence  
CC FEATURE:  
CC OTHER INFORMATION: Description of Artificial Sequence: Fusion  
CC OTHER INFORMATION: construct of human CD39  
CC SEQUENCE 463 AA; 52621 MW; 1178782 CN;  
Query Match 100.0%; Score 3275; DB 1; Length 463;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 25 TONKALPENKYGIVLDAGSSHTSLYIKWPAKENDTGVAHVQVEBCRVKPGISKPVOK 84  
| | | | |  
Qy 1 TONKALPENKYGIVLDAGSSHTSLYIKWPAKENDTGVAHVQVEBCRVKPGISKPVOK 60  
| | | | |  
Db 85 VNEIGIYLTDCMERAREVIPRSOHOETPYVLTAGTMRLLRMESEELADRVLDYVERSLS 144  
| | | | |  
Qy 61 VNEIGIYLTDCMERAREVIPRSOHOETPYVLTAGTMRLLRMESEELADRVLDYVERSLS 120  
| | | | |  
Db 145 NYPFDFOGARIIITGOEGAGYGMTITNYLLGKFSQKTRWFSIVPYETNNQETFGALDLGA 204  
| | | | |  
Qy 121 NYPFDFOGARIIITGOEGAGYGMTITNYLLGKFSQKTRWFSIVPYETNNQETFGALDLGA 180  
| | | | |  
Db 205 STQVTFVPQNOTIESPDNALQFRLYGKDYNVYTHSFLCYGKDQALMOKLAKIOVASNEI 264  
| | | | |  
Qy 181 STQVTFVPQNOTIESPDNALQFRLYGKDYNVYTHSFLCYGKDQALMOKLAKIOVASNEI 240  
| | | | |



Db	265	LADPCFHGKXKVVNVDLYTCTPCKREMTLPQOQEEIOIGNYQOCHOSILEFNTSY	324
Oy	241	LADPCFHGKXKVVNVDLYTCTPCKREMTLPQOQEEIOIGNYQOCHOSILEFNTSY	300
Db	325	CPYSCAFNGJFLPPLQDGFASFVFMKFLMLTSEKYSOEKVYTMKMKFCAQPEEI	384
Oy	301	CPYSCAFNGJFLPPLQDGFASFVFMKFLMLTSEKYSOEKVYTMKMKFCAQPEEI	360
Db	385	KTSYAGVKEKYLSECESGTYIISLLQGYHFTADSMEHIFITKIQGSDAGMTLGMLN	444
Oy	361	KTSYAGVKEKYLSECESGTYIISLLQGYHFTADSMEHIFITKIQGSDAGMTLGMLN	420
Db	445	LTNMIPAEQPLSTPLSHST	463
Oy	421	LTNMIPAEQPLSTPLSHST	439
RESULT	4	STANDARD:	PRT: 464 AA.
ID	PCF-US99-22955-27		
AC	xxxxxx		
DT			
XX			
DE	Sequence 27, Application PC/TUS9922955		
XX			
CC	Sequence 27, Application PC/TUS9922955		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Mallszewski, Charles R.		
CC	APPLICANT: Gayle III, Richard B.		
CC	APPLICANT: Price, Virginia L.		
CC	APPLICANT: Grice, Steven D.		
CC	APPLICANT: Immunex Corporation		
CC	TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment		
CC	FILE REFERENCE: 2879-WO		
CC	CURRENT APPLICATION NUMBER: PCF/US99/22955		
CC	CURRENT FILING DATE: 1999-10-13		
CC	EARLIER APPLICATION NUMBER: US 60/104,585		
CC	EARLIER FILING DATE: 1998-10-16		
CC	EARLIER APPLICATION NUMBER: US 60/107,466		
CC	EARLIER FILING DATE: 1998-11-06		
CC	EARLIER APPLICATION NUMBER: US 60/149,010		
CC	EARLIER FILING DATE: 1999-08-13		
CC	NUMBER OF SEQ ID NOS: 31		
CC	SOFTWARE: PatentIn Ver. 2.0		
CC	SEQ ID NO 27		
CC	LENGTH: 464		
CC	TYPE: PRT		
CC	ORGANISM: Artificial Sequence		
CC	FEATURE:		
CC	OTHER INFORMATION: Description of Artificial Sequence: Fusion		
CC	OTHER INFORMATION: construct of human CD39		
CC	SEQUENCE 464 AA: 52754 MW; 1182698 CN;		
SO			
Query Match	100.0%;	Score 3275;	DB 1; Length 464;
Best Local Similarity	100.0%;	Pred. No. 0.00e+00;	
Matches 439;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Db	26	TONKALPENNVKGYVLADGSSHTSLYIKWPAEKENDTGVVHOVEBCRVGPGISKRVOK	85
Oy	1	TONKALPENNVKGYVLADGSSHTSLYIKWPAEKENDTGVVHOVEBCRVGPGISKRVOK	60
Db	86	VNEIGIYITLDCMERAREVTPRSQOHEPVPYIGATAGMKMLRMESEELADRYLDVVERSL	145
Oy	61	VNEIGIYITLDCMERAREVTPRSQOHEPVPYIGATAGMKMLRMESEELADRYLDVVERSL	120
Db	146	NYPFDFGACARIITGOEGAGVMTITNYLLGKFSQKTRWFSIVPYETNNQETFGALDIGA	205
Oy	121	NYPFDFGACARIITGOEGAGVMTITNYLLGKFSQKTRWFSIVPYETNNQETFGALDIGA	180
Db	206	STQVTFVFNQNTIESPDNALQFRILYKGDYNYVTHSFLCYGKDQALNKLAKDIOVASNEI	265
Oy	181	STQVTFVFNQNTIESPDNALQFRILYKGDYNYVTHSFLCYGKDQALNKLAKDIOVASNEI	240

Dd	266	LADPCHFGYKRVVNSDLYTPTCTKRREMLLPFOQIEIGICGVQCCHSILEFTSTY	325
Oy	241	LADPCHFGYKRVVNSDLYTPTCTKRREMLPFPOQIEIGICGVQCCHSILEFTSTY	300
Dd	326	CPSYSCANGFIFLPILODFGAFSAFYVMKFNLUTSEKVSQEKVTMMKKFCQAOPBEI	385
Oy	301	CPSYSCANGFIFLPILODFGAFSAFYVMKFNLUTSEKVSQEKVTMMKKFCQAOPBEI	360
Dd	386	KTSYAGVEKEKLSEXCFSGYILSLLLQGYPHTADSMEHIFIGKIQGSAGMTLGMYLN	445
Oy	361	KTSYAGVEKEKLSEXCFSGYILSLLLQGYPHTADSMEHIFIGKIQGSAGMTLGMYLN	420
Dd	446	LTNMTPAEOPPLSTPLSHST 464	
Oy	421	LTNMTPAEOPPLSTPLSHST 439	
 RESULT            5            STANDARD;         PRT:    473 AA.			
ID	PCT-US99-22955-29		
XX	xxxxxx		
XX			
DT			
XX			
DE	Sequence 29, Application PC/TUS9922955		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Maliszewski, Charles R.		
CC	APPLICANT: Gayle III, Richard B.		
CC	APPLICANT: Price, Virginia L.		
CC	APPLICANT: Gimpel, Steven D.		
CC	APPLICANT: Immunex Corporation		
CC	TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment		
CC	FILE REFERENCE: 2879-WO		
CC	CURRENT APPLICATION NUMBER: PCT/US99/22955		
CC	CURRENT FILING DATE: 1999-10-13		
CC	EARLIER APPLICATION NUMBER: US 60/104,585		
CC	EARLIER FILING DATE: 1998-10-16		
CC	EARLIER APPLICATION NUMBER: US 60/107,466		
CC	EARLIER FILING DATE: 1998-11-06		
CC	EARLIER APPLICATION NUMBER: US 60/149,010		
CC	EARLIER FILING DATE: 1999-08-13		
CC	NUMBER OF SEQ ID NOS: 31		
CC	SOFTWARE: PatentIn Ver. 2.0		
CC	SEQ ID NO 29		
CC	LENGTH: 473		
CC	TYPE: PRt		
CC	ORGANISM: Artificial Sequence		
CC	FEATURE:		
CC	OTHER INFORMATION: Description of Artificial Sequence: Fusion		
CC	OTHER INFORMATION: construct of human CD39		
SQ	SEQUENCE    473 AA: 53745 MW; 1229492 CN;		
 Query Match            100.0%; Score 3275; DB 1; Length 473: Best Local Similarity 100.0%: Pred. NO. 0.00e+00; Matches    439; Conservative    0; Mismatches    0; Indels    0; Gaps    0;			
Dd	35	TGNKALPENVKXGIVLDAGSSHTSLTIYIKWPAEKENDTGVVHOYECECRVKGPGISKRFVK	94
Oy	1	TGNKALPENVKXGIVLDAGSSHTSLTIYIKWPAEKENDTGVVHOYECRVRKPGISKRFVK	60
Dd	95	VNEITGIYLTDCMERAREVIRPSQHOFETPVYVGATPAGMKLMRMESEELADRYLDIVERSLS	154
Oy	61	VNEITGIYLTDCMERAREVIRPSQHOFETPVYVGATPAGMKLMRMESEELADRYLDIVERSLS	120
Dd	155	NYPFPQGARITTGEEBAVGMITINTYLGLGFQSOTKTFISIVPYETNNQETFGALDGA	214
Oy	121	NYPFPQGARITTGEEBAVGMITINTYLGLGFQSOTKTFISIVPETNNQETFGALDGA	180
Dd	215	STGVTFPNQNTIESPDNALFRLYGKDYNVYTHSFCLCYGKQDALMOKLAKDIVASNET	274

QY	181	STQYTFVPQNTIESPDNALQFRILXGNDYNTYHSPFLCYGQDQALMQKLAKNDIQVASNEI	260
Db	275	LRDPCFHFQYKKVVNVSDLYTPTCTKREEMTLPRQFQELQIGNYQQCHOSILELNTST	334
QY	241	LRRDPCFHFQYKKVVNVSDLYTPTCTKREEMTLPRQFQELQIGNYQQCHOSILELNTSTY	300
Db	335	CPSQCAFNGLFLPRLQDQFCAFSAFYVMKFLMLTSEKYSQEKVTEMMKFFCAQPIEEI	394
QY	301	CPSQCAFNGLFLPRLQDQFCAFSAFYVMKFLMLTSEKYSQEKVTEMMKFFCAQPIEEI	360
Db	395	KTSYAGVAKELYSEKCFESGTYILSLLOGYHFTADSMENHIFICKIQGSDAGWTLGYMLN	454
QY	361	KTSYAGVAKELYSEKCFESGTYILSLLOGYHFTADSMENHIFICKIQGSDAGWTLGYMLN	420
Db	455	LTNMTPAEQPLSTPLSHST	473
QY	421	LTNMTPAEQPLSTPLSHST	439

RESULT	6	STANDARD;	PRT;	474	AA
ID	PCT-US99-22955-28				

Sequence 28, Application PC/TUS9922955

```

CC Sequence 28, Application PC/TUS9922255
CC GENERAL INFORMATION:
CC APPLICANT: Maliszewski, Charles R.
CC APPLICANT: Gayle III, Richard B.
CC APPLICANT: Price, Virginia L.
CC APPLICANT: Gimpel, Steven D.
CC APPLICANT: Immunex Corporation
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
CC FILE REFERENCE: 2879-WO
CC CURRENT APPLICATION NUMBER: PCT/TUS99/22955
CC CURRENT FILING DATE: 1999-10-13
CC EARLIER APPLICATION NUMBER: US 60/104,585
CC EARLIER FILING DATE: 1998-10-16
CC EARLIER APPLICATION NUMBER: US 60/107,466
CC EARLIER FILING DATE: 1998-11-06
CC EARLIER APPLICATION NUMBER: US 60/149,010
CC EARLIER FILING DATE: 1999-08-13
CC NUMBER OF SEQ. ID NOS: 31
CC SOFTWARE: PatentIn Ver. 2.0
CC SEQ ID NO 28
CC
CC LENGTH: 474
CC TYPE: PRT
CC ORGANISM: Artificial Sequence
CC FEATURE:
CC OTHER INFORMATION: Description of Artificial Sequence: Fusion
CC OTHER INFORMATION: construct of human CD39
CC SEQUENCE 474 AA: 53816 MW: 1234400 CN:

```

QY	181	STQYTFVFNQNTIEBPDNALDFRLTKGDYINVYTHSFLCYGDAQMLMKLANKDQVANSNEI	240
Db	276	LRDPCFHPGKRYKVVNSDLVYTPCTKRFEMTLPRQDFEIGICGNYQOCHSILEFNSTY	335
QY	241	LRDPEFHGKRYKVVNSDLVYTPCTKRFEMTLPRQDFEIGICGNYQOCHSILEFNSTY	300
Db	336	CPYSQACFNAGJFLPRLQDGFCAFSAFYVVMKFLMLTSEKVSQEKVYTMKMKFCQAQWEEI	395
QY	301	CPYSQACFNAGJFLPRLQDGFCAFSAFYVVMKFLMLTSEKVSQEKVYTMKMKFCQAQWEEI	360
Db	396	KTSYGVGVEKYLISEKCFESGTYILSLLOGGYHFTADSMEHFIIGKIOGSDAGWTICMLN	455
QY	361	KTSYGVGVEKYLISEKCFESGTYILSLLOGGYHFTADSMEHFIIGKIOGSDAGWTICMLN	420
Db	456	LTNMIPAEQPLSTPLSHST	474
QY	421	LTNMIPAEQPLSTPLSHST	439

RESULT	7
ID	PCT-US99-22955-3
STANDARD;	PRT;
	476 AA.

Sequence 3, Application PC/TUS9922955

CC	Sequence 3, Application PC/US99922955
CC	GENERAL INFORMATION:
CC	APPLICANT: Maliszewski, Charles R.
CC	APPLICANT: Gayle III, Richard B.
CC	APPLICANT: Price, Virginia L.
CC	APPLICANT: Gimpel, Steven D.
CC	APPLICANT: Immunex Corporation
CC	TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
CC	FILE REFERENCE: 2879-WO
CC	CURRENT APPLICATION NUMBER: PCT/US99/22955
CC	CURRENT FILING DATE: 1999-10-13
CC	EARLIER APPLICATION NUMBER: US 60/104,585
CC	EARLIER FILING DATE: 1998-10-16
CC	EARLIER APPLICATION NUMBER: US 60/107,466
CC	EARLIER FILING DATE: 1998-11-06
CC	EARLIER APPLICATION NUMBER: US 60/149,010
CC	EARLIER FILING DATE: 1999-08-13
CC	NUMBER OF SEQ ID NOS: 31
CC	SOFTWARE: PatentIn Ver. 2.0
CC	SEQ ID NO 3
CC	LENGTH: 476
CC	TYPE: PR1
CC	ORGANISM: Artificial Sequence
CC	FEATURE:
CC	OTHER INFORMATION: Description of Artificial Sequence: Fusion
CC	OTHER INFORMATION: construct of human CD39
CC	SEQUENCE 476 AA: 54177 MW: 1244820 CN:

Db 40 TONALPENKYGIVLADGSSHTSLYTYKKPAEKENDTGVHVOEBCRVKPGISLFFVOK 99  
 |||||  
 Qy 1 TONALPENKYGIVLADGSSHTSLYTYKKPAEKENDTGVHVOEBCRVKPGISLFFVOK 60  
 |||||  
 Db 100 VNEIGIYLJDCMERAREVJPRSOHOETPVYLGATAGKRLLRMESEELADRVLDVERSTLS 159  
 |||||  
 Qy 61 VNEIGIYLJDCMERAREVJPRSOHOETPVYLGATAGKRLLRMESEELADRVLDVERSTLS 120  
 |||||  
 Db 160 NYPPDFOGARITIQOECAVGMITINLLKFKSOKTWFSIVEYETJNNOTFCALDLGGA 219  
 |||||  
 Qy 121 NYPPDFOGARITIQOECAVGMITINLLKFKSOKTWFSIVEYETJNNOTFCALDLGGA 180  
 |||||  
 Db 220 STOVTEVPQNTIESPDNALQFLYGDVNVYTHSEFLCYGKDQALMOKLAKDIQVANSNEI 279  
 |||||

Db	49	TONALEENKYGIVLADGSSHTSLYTKKMPAEKENDTGVHVOEECRVPGISFVK	108
Qy	1	TONALEENKYGIVLADGSSHTSLYTKKMPAEKENDTGVHVOEECRVPGISFVK	60
Db	109	VNEIGIVLTCOMERAREVIRPSQHOETPVYLGATAGMRLLRMESEELADRVLDVERSSLS	168
Qy	61	VNEIGIVLTCOMERAREVIRPSQHOETPVYLGATAGMRLLRMESEELADRVLDVERSSLS	120
Db	169	NYPDFOGARITIGOECAAGWITINLLKFSOKTRMESIVLEYETINNOETFGALDLGA	228
Qy	121	NYPDFOGARITIGOECAAGWITINLLKFSOKTRMESIVLEYETINNOETFGALDLGA	180
Db	229	STOVTEVPONOTIESPNNALOFRLYGKDYNVYHSHFCYCKDQALOMOKLAKDIQVANSNEI	288
Qy	181	STOVTEVPONOTIESPNNALOFRLYGKDYNVYHSHFCYCKDQALOMOKLAKDIQVANSNEI	240

Db	289	LRCEFHGGYKKVNVNSLYTKPTCKRREMTLPQOEIIGISGYOCHOSIILETNTSY	348
Qy	241	LRDPFHGGYKKVNVNSLYTKPTCKRREMTLPQOEIIGISGYOCHOSIILETNTSY	300
Db	349	CPYSQCAFNGLFPLPLQDGFASAFYFVVKFTLNLTSSEKYSOEKVTMMKPCAQWEIEI	408
Qy	301	CPYSQCAFNGLFPLPLQDGFASAFYFVVKFTLNLTSSEKYSOEKVTMMKPCAQWEIEI	360
Db	409	KTSIAGVKEKLTSEYCSGGVITLSLLQGYHFTADSENIHFIGKIQGSDAGMTIGYMLN	468
Qy	361	KTSIAGVKEKLTSEYCSGGVITLSLLQGYHFTADSENIHFIGKIQGSDAGMTIGYMLN	420
Db	469	LTNMIAPAOPLSTPLSHT	487
Qy	421	LTNMIAPAOPLSTPLSHT	439

	RESULT	10			PRT;	510 AA.
ID	US-08-	930-921-1	STANDARD;			
XX						
AC	xxxxxx					
XX						
DT						
DE	Sequence 1,	Application US/08930921B				
CC						
CC	Sequence 1,	Application US/08930921B				
CC	GENERAL INFORMATION:					
CC	APPLICANT:	BEAUDOIN, Adrien R.				
CC	APPLICANT:	SEVIGNY, Jean				
CC	TITLE OF INVENTION:	ATP-DIPHOSPHOHYDROLASES, PROCESS OF PURIFICATION				
CC	TITLE OF INVENTION:	THEREOF AND PROCESS OF PRODUCING THEREOF BY RECOMBINANT				
CC	TITLE OF INVENTION:	TECHNOLOGY				
CC	FILE REFERENCE:	ATP-DIPHOSPHOHYDROLASES-BEADOIN				
CC	CURRENT APPLICATION NUMBER:	US/08/930.921B				
CC	CURRENT FILING DATE:	1996-01-02				
CC	EARLIER APPLICATION NUMBER:	PCF/CA96/00223				
CC	EARLIER FILING DATE:	1996-04-10				
CC	NUMBER OF SEQ ID NOS:	8				
CC	SOFTWARE:	PatentIn Ver. 2.0				
CC	SEQ ID NO 1					
CC	LENGTH:	510				
CC	TYPE:	PRT				
CC	ORGANISM:	Unknown				
CC	FEATURE:					
CC	OTHER INFORMATION:	Description of Unknown Organism: unknown				
Q0	SEQUENCE	510 AA; 57964 MW; 1469467 CN;				

	Query Match	100.0%	Score 3275;	DB 14;	Length 510;
	Best Local Similarity	100.0%	Pred. No. 0.00e+00;		
	Matches 439;	Conservative	0;	Mismatches 0;	Indels 0;
				Gaps	0;
Db	38	TQNALPENNYGYIVLDAGSSHTSLTYIKYKPAEKNENTGVVHOVEECRYVPGISKRVOK	97		
Qy	1	TQNALPENNYGYIVLDAGSSHTSLTYIKYKPAEKNENTGVVHOVEECRYVPGISKRVOK	60		
Db	98	VNEIGIYLTDCMERAREVYIPRSQHOETPVYLGATAGRLLRMESEELADAVLYDVERKSLS	157		
Qy	61	VNEIGIYLTDCMERAREVYIPRSQHOETPVYLGATAGRLLRMESEELADAVLYDVERKSLS	120		
Db	158	NYPDFOGARITITGOEBGANGTITNTLLGKFSOKTFWSTIVPEYETINNOETFGALDIGA	217		
Qy	121	NYPDFOGARITITGOEBGANGTITNTLLGKFSOKTFWSTIVPEYETINNOETFGALDIGA	180		
Db	218	STQVTFVPQNTIESPNALQFRLYGDNVYTHSFICYKQDALMOKLAKDQVANSNEI	277		
Qy	181	STQVTFVPQNTIESPNALQFRLYGDNVYTHSFICYKQDALMOKLAKDQVANSNEI	240		
Db	278	LRDQCFPHGYKVVVNSDLVTKPCTKRFEEMTLPPQOEIIGIGVQOCHOSILEFNSTYS	337		
Qy	241	LRDQCFPHGYKVVVNSDLVTKPCTKRFEEMTLPPQOEIIGIGVQOCHOSILEFNSTYS	300		
Db	338	CPSQACANGITLPLLOGDGFASAFYVVKFLNTLTSEKVSQEKVTEMMKPCAOFWEEI	397		

QY 301 CYYSCAENGITLPLPLQDPCGFSAFYVMFLNITSKVSQEVNTEMMKKFCAQPPBEI 360

Db 398 KTSYAGVAREKYLSEYCFSGTYILSLLOGYHPTADSWEHIFHTKIOGSDAGWTLYGMLN 457

QY 361 KTSYAGVAREKYLSEYCFSGTYILSLLOGYHPTADSWEHIFHTKIOGSDAGWTLYGMLN 420

Db 458 LTNMTPAEOPPLSTPLSHST 476

QY 421 LTNMTPAEOPPLSTPLSHST 439

RESULT	11	STANDARD	PRT	510 AA.
ID	PCT-US99-22955-2			
AC	xxxxxx			
DT				
XX				
DE	Sequence 2, Application PC/TUS9922955			
XX				
CC	Sequence 2, Application PC/TUS9922955			
CC	GENERAL INFORMATION:			
CC	APPLICANT: Maliszewski, Charles R.			
CC	APPLICANT: Gayle III, Richard B.			
CC	APPLICANT: Price, Virginia L.			
CC	APPLICANT: Gimpel, Steven D.			
CC	APPLICANT: Immunex Corporation			
CC	TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment			
CC	FILE REFERENCE: 2879-WO			
CC	CURRENT APPLICATION NUMBER: PCT/US99/22955			
CC	CURRENT FILING DATE: 1999-10-13			
CC	EARLIER APPLICATION NUMBER: US 60/104,585			
CC	EARLIER FILING DATE: 1998-10-16			
CC	EARLIER APPLICATION NUMBER: US 60/107,466			
CC	EARLIER FILING DATE: 1998-11-06			
CC	EARLIER APPLICATION NUMBER: US 60/149,010			
CC	EARLIER FILING DATE: 1999-08-13			
CC	NUMBER OF SEQ ID NOS: 31			
CC	SOFTWARE: PatentIn Ver. 2.0			
CC	SEQ ID NO 2			
CC	LENGTH: 510			
CC	TYPE: PRT			
CC	ORGANISM: Homo sapiens			
CC	SEQUENCE 510 AA: 57964 MW: 1469467 CN:			

	Query Match	100.0%	Score 3275	DB 1	Length 510
	Best Local Similarity 100.0%		Pred. No. 0.00e+00		
	Matches 439	Conservative	0	Mismatches 0	Indels 0
				Gaps 0	
Db	38	TQNKLPENVKYGILADGSSHTSLYIKKPAEKENDTGVHVQEBCRVKPGISKFKVOK	97		
Qy	1	TQNKLPENVKYGILADGSSHTSLYIKKPAEKENDTGVHVQEBCRVKPGISKFKVOK	60		
Db	98	VNEIGIYLTDCMERAREVIPRSQHOETPPVYLGAATAGRLLRMESEELADRYLDVERST	157		
Qy	61	VNEIGIYLTDCMERAREVIPRSQHOETPPVYLGAATAGRLLRMESEELADRYLDVERST	120		
Db	158	NYPDFOGARIIITGOEBAGVMTINTILNLFKFSQKTFWSTSVPEETNNQETFGALDIGA	217		
Qy	121	NYPDFOGARIIITGOEBAGVMTINTILNLFKFSQKTFWSTSVPEETNNQETFGALDIGA	180		
Db	218	STQVTFVQNTIETSPDIALQFLRYGDYNYTHSFLCYCKDQALMOKLAKDIQVANSNEI	277		
Qy	181	STQVTFVQNTIETSPDIALQFLRYGDYNYTHSFLCYCKDQALMOKLAKDIQVANSNEI	240		
Db	278	LRDPCFHHGKYKVVNVNSLYKTPCTKRFEEMTLPRQCEIIGIGNYQOCHOSILELNTSY	337		
Qy	241	LRDPCFHHGKYKVVNVNSLYKTPCTKRFEEMTLPRQCEIIGIGNYQOCHOSILELNTSY	300		
Db	338	CPYISQCAFNGLFPLPDGDFGASAFYFVAKFLNLISEKYSQEKVTEEMMKFCAQWEET	397		
Qy	301	CPYISQCAFNGLFPLPDGDFGASAFYFVAKFLNLISEKYSQEKVTEEMMKFCAQWEET	360		

Db	398	KTSAVGEKTLSECFEGTIIISLLQGYFTADSMEHIFGKIOGSDAGMTLGYMLN	457
Qy	361	KTSAVGEKTLSECFEGTIIISLLQGYFTADSMEHIFGKIOGSDAGMTLGYMLN	420
Db	458	LTNMIPAEQPLSTPLSHST	476
Qy	421	LTNMIPAEQPLSTPLSHST	439
RESULT	12		
ID	US-08-701-460-2	STANDARD:	PRT: 510 AA.
XX	xxxxxx		
DT			
XX			
XX			
DE	Sequence 2, Application US/08701460		
XX			
CC	Sequence 2, Application US/08701460		
CC	GENERAL INFORMATION:		
CC	APPLICANT: MARCUS, AARON J.		
CC	APPLICANT: MALISZEWSKI, CHARLES R.		
CC	APPLICANT: GAYLE, RICHARD B.		
CC	TITLE OF INVENTION: METHODS OF REGULATING HEMOSTASIS AND		
CC	TITLE OF INVENTION: THROMBOSIS, AND COMPOUNDS USEFUL THEREFOR		
CC	NUMBER OF SEQUENCES: 16		
CC	CORRESPONDENCE ADDRESSES:		
CC	ADDRESSEE: Immunex Corporation		
CC	STREET: 51 University Street		
CC	City: Seattle		
CC	STATE: WA		
CC	COUNTRY: USA		
CC	ZIP: 98101		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: Apple Power Macintosh 7200/30		
CC	OPERATING SYSTEM: Apple Operating System 7.5.3		
CC	SOFTWARE: Microsoft Word 6.0.1 for Power Macintosh		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/701,460		
CC	FILING DATE: 22 AUGUST 1996		
CC	CLASSIFICATION: 514		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: Perkins, Patricia Anne		
CC	REGISTRATION NUMBER: 34,693		
CC	REFERENCE/DOCKET NUMBER: 2808		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: 206-587-0430		
CC	INFORMATION FOR SEQ ID NO: 2:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 510 amino acids		
CC	TYPE: amino acid		
CC	TOPOLOGY: linear		
CC	MOLECULE TYPE: protein		
CC	SEQUENCE	510 AA; 57964 MW; 1469467 CN;	
Qy	Query Match	100.0%; Score 3275; DB 12; Length 510;	
Db	Best Local Similarity	100.0%; Pred. No. 0.00e+00;	
Qy	Matches 439; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Db	38	TQNALPENNVKYGIVLDAGSSHTSLYIKWPAEKENDTGVAHVQEECRVYGPGISKEFVKQ	97
Qy	1	TQNALPENNVKYGIVLDAGSSHTSLYIKWPAEKENDTGVAHVQEECRVYGPGISKEFVKQ	60
Db	98	VNEIGIYITDCMEKAREVYIPRSQHQETPVYLGATAGKRLRLMSESEELADRYLDIVERSLS	157
Qy	61	VNEIGIYITDCMEKAREVYIPRSQHQETPVYLGATAGKRLRLMSESEELADRYLDIVERSLS	120
Db	158	NYPDPGQARITLQOEAGAGWITINTLLKFSQKTRWESIYPERTNNOETFGALDIGA	217
Qy	121	NYPDPGQARITLQOEAGAGWITINTLLKFSQKTRWESIYPERTNNOETFGALDIGA	180

D	b	218	STQTFVPOANTTISPDNALQFRLYGKDYNYTHSFCLCYGKDALWOKLAKDIQVASNEI	277
O	y	181	STQTFVPOANTTISPDNALQFRLYGKDYNYTHSFCLCYGKDALWOKLAKDIQVASNEI	240
D	b	278	LRDPCFHFGYKKVVNVSDLYKTPTKRFEEMTLPPQOFEIQIGNYOQCHOSILELFNTSY	337
O	y	241	LRDPCFHFGYKKVVNVSDLYKTPTKRFEEMTLPPQOFEIQIGNYOQCHOSILELFNTSY	300
D	b	338	CPSYCQAFNGIFLEPPLQDGFAGFSAPFYFMKFLNLJSEKVSQEKYTEMMKKPCAQPWEET	397
O	y	301	CPSYCQAFNGIFLEPPLQDGFAGFSAPFYFMKFLNLJSEKVSQEKYTEMMKKPCAQPWEET	360
D	b	398	KTSVAGYEKELSYCSCGYITISLLQGYNFTADSNHEHIFTGIQOGSAGMTGLGYMLN	457
O	y	361	KTSVAGYEKELSYCSCGYITISLLQGYNFTADSNHEHIFTGIQOGSAGMTGLGYMLN	420
D	b	458	LTNMIPAEOPLSTPLSHST	476
O	y	421	LTNMIPAEOPLSTPLSHST	439
RESULT 13 STANDARD: PRT; 510 AA.				
ID	XX	US-09-374-586-1	SEQUENCE 1, Application US/09374586	
AC	XXXXXX			
DT				
CC	CC	Sequence 1, Application US/09374586		
CC	CC	GENERAL INFORMATION:		
CC	CC	APPLICANT: Plinsky, David J.		
CC	CC	TITLE OF INVENTION: CD39/ECTO-ADPASE AS A TREATMENT FOR THROMBOTIC AND		
CC	CC	TITLE OF INVENTION: ISCHEMIC DISORDERS		
CC	CC	FILE REFERENCE: 0575/59167		
CC	CC	CURRENT APPLICATION NUMBER: US/09/374, 586		
CC	CC	CURRENT FILING DATE: 1999-08-13		
CC	CC	NUMBER OF SEQ ID NOS: 2		
CC	CC	SOFTWARE: PatentIn Ver. 2.1		
CC	CC	SEQ ID NO 1		
CC	CC	LENGTH: 510		
CC	CC	TYPE: PRT		
CC	CC	ORGANISM: HOMO-SAPIEN		
SQ	SQ	SEQUENCE 510 AA; 57964 MW; 1469467 CN;		
Query Match 100.0%; Score 3275; DB 25; Length 510;				
Best Local Similarity 100.0%; Pred. No. 0.00e+00;				
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
D	b	38	TQNKLDPENVXYGIVLDAGSSHTSLYTKKPAEKENDTGVAHVQEBCRVKGPISKRFVK	97
O	y	1	TQNKLDPENVXYGIVLDAGSSHTSLYTKKPAEKENDTGVAHVQEBCRVKGPISKRFVK	60
D	b	98	VNEIGIYITDCMEAREVIIPRSQHQTPEVYLGAAGRLLRMESELADRYLDIVERSLS	157
O	y	61	VNEIGIYITDCMEAREVIIPRSQHQTPEVYLGAAGRLLRMESELADRYLDIVERSLS	120
D	b	158	NYPDPGARITIQOEGSAIGMITINLLDKFSQKTFWESIYVENNQOTFCALDLGA	217
O	y	121	NYPDPGARITIQOEGSAIGMITINLLDKFSQKTFWESIYVENNQOTFCALDLGA	180
D	b	218	STQTFVPONOTISSPNALQFRLYGKDYVNYTHSFCLCYGKDOLWOKLAKDIQVASNEI	277
O	y	181	STQTFVPONOTISSPNALQFRLYGKDYVNYTHSFCLCYGKDOLWOKLAKDIQVASNEI	240
D	b	278	LARDPCFHFGYKKVVNVSDLYKTPTKRFEEMTLPPQOFEIQIGNYOQCHOSILELFNTSY	337
O	y	241	LARDPCFHFGYKKVVNVSDLYKTPTKRFEEMTLPPQOFEIQIGNYOQCHOSILELFNTSY	300
D	b	338	CPSYCQAFNGIFLEPPLQDGFAGFSAPFYFMKFLNLJSEKVSQEKYTEMMKKPCAQPWEET	397
O	y	301	CPSYCQAFNGIFLEPPLQDGFAGFSAPFYFMKFLNLJSEKVSQEKYTEMMKKPCAQPWEET	360

Dd	398	KTSVAGVEKTLSEKCEFGTYLLSLGLQGYHPTADSWGHIFPICKIOGSDAGMTLGMNL	457
Oy	361		420
Dd	458	LTMNIPAEOPLPSTPLSHST	476
Oy	421		439
Oy	421	LTMNIPAEOPLPSTPLSHST	439
RESULT	14	STANDARD:	PRT; 476 AA.
ID	PCT-US99-22955-4		
XX	AC	xxxxxx	
Dt	Sequence 4, Application PC/TUS9922955		
XX	DE		
CC	Sequence 4, Application PC/TUS9922955		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Maliszewski, Charles R.		
CC	APPLICANT: Gayle III, Richard B.		
CC	APPLICANT: Price, Virginia L.		
CC	APPLICANT: Gipmel, Steven D.		
CC	APPLICANT: Immunex Corporation		
CC	TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment		
CC	FILE REFERENCE: 2879-WO		
CC	CURRENT APPLICATION NUMBER: PCT/US99/22955		
CC	CURRENT FILING DATE: 1999-10-13		
CC	EARLIER APPLICATION NUMBER: US 60/104,585		
CC	EARLIER FILING DATE: 1998-10-16		
CC	EARLIER APPLICATION NUMBER: US 60/107,466		
CC	EARLIER FILING DATE: 1998-11-06		
CC	EARLIER APPLICATION NUMBER: US 60/149,010		
CC	EARLIER FILING DATE: 1999-08-13		
CC	NUMBER OF SEQ ID NOS: 31		
CC	SOFTWARE: PatentIn Ver. 2.0		
CC	SEQ ID NO 4		
CC	LENGTH: 476		
CC	TYPE: PRT		
CC	ORGANISM: Artificial Sequence		
CC	FEATURE:		
CC	OTHER INFORMATION: Description of Artificial Sequence: Fusion		
CC	OTHER INFORMATION: construct of human CD39		
CC	FEATURE:		
CC	NAME/KEY: VARIANT		
CC	LOCATION: (39)		
CC	OTHER INFORMATION: Any amino acid, preferably Cys or Ser		
SQ	SEQUENCE 476 AA; 54078 MW; 1246383 CN;		
Query Match	97.9%; Score 3206; DB 1; Length 476;		
Best Local Similarity	100.0%; Pred. No. 0.00e+00;		
Matches 426; Conservative	0; Mismatches 0; Indels 0; Gaps 0		
Dd	49	YGIIVDAGSSHTSLIYKPMPAEKENDTGVAQVEBCIRKPGSGISKFYQKVMEIGIYLTDG	108
Oy	12	YGIIVDASSHTSLIYKPMPAEKENDTGVAQVEBCIRKPGSGISKFYQKVMEIGIYLTDG	71
Dd	109	MERAREVIPRSOHETPVYLGATAGMARLLRMESEELADRVLDDVVERISLNTPEDFOGARI	168
Oy	72	MERAREVIPRSOHETPVYLGATAGMARLLRMESEELADRVLDDVVERISLNTPEDFOGARI	131
Dd	169	ITGOEGAVYGTTINYYLKGFSQKRMSIYPIETNNGETGALDIGASQVTFVPONQ	228
Oy	132	ITGOEGAVYGMTTINYLLKGFSQKRMSIYPIETNNGETGALDIGASQVTFVPONQ	191
Dd	229	TIESPDNALQRLKYGDYNYTTHSEFLCYGKDQALMOKLAKDIQVASNEILLDPGFHHGYK	288
Oy	192	TIESPDNALQRLKYGDYNYTTHSEFLCYGKDQALMOKLAKDIQVASNEILLDPGFHHGYK	251
Dd	289	KVVNVSDLYKTPCKRFEMTLRPQQFELIQIGNTVOOQHSLLELFNTSYCPYSOCAPENG	348

```

0Y 252 KVVNNSDLKPKPCIKREPMTELPFOQFEIQQIGNQQOCHQSHLLEFNINSYCPISQCAFNGI 311
Db 349 FLPLPQDGFASAFYFWKFLNLTSEKVSQEKYTEMMKKFCQAOPWBEIKTSYAGVAREKY 408
0Y 312 FLPLPQDGFASAFYFWKFLNLTSEKVSQEKYTEMMKKFCQAOPWBEIKTSYAGVAREKY 371
Db 409 LSEYFSGSTYIISLLQGYHTFADSWEHHTIGTKIQGSDAGWTIGYMLNLTNMTIAPAOPL 468
0Y 372 LSEYFSGSTYIISLLQGYHTFADSWEHHTIGTKIQGSDAGWTIGYMLNLTNMTIAPAOPL 431
Db 469 STPLSHST 476
0Y 432 STPLSHST 439

RESULT 15
ID US-08-701-460-4 STANDARD; PRT; 510 AA.
XX
XX
XX
XX
XX
DE
XX
XX
Sequence 4, Application US/08701460
CC
CC Sequence 4, Application US/08701460
CC GENERAL INFORMATION:
CC APPLICANT: MARCUS, AARON J.
CC APPLICANT: MALISZEWSKI, CHARLES R.
CC APPLICANT: GAYLE, RICHARD B.
CC TITLE OF INVENTION: METHODS OF REGULATING HEMOSTASIS AND
CC TITLE OF INVENTION: THROMBOSIS, AND COMPOUNDS USEFUL THEREFOR
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Immunex Corporation
CC STREET: 51 University Street
CC CITY: Seattle
CC STATE: WA
CC COUNTRY: USA
CC ZIP: 98101
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: Apple Power Macintosh 7200/90
CC OPERATING SYSTEM: Apple Operating System 7.5.3
CC SOFTWARE: Microsoft Word 6.0.1 for Power Macintosh
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/701,460
CC FILING DATE: 22 AUGUST 1996
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Perkins, Patricia Anne
CC REGISTRATION NUMBER: 34,693
CC REFERENCE/DOCKET NUMBER: 2808
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 206-587-0430
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 510 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 510 AA; 57205 MW; 1483126 CN;

Query Match 80.8%; Score 2645; DB 12; Length 510;
Best Local Similarity 76.6%; Pred. No. 4,51e-280;
Matches 339; Conservative 57; Mismatches 42; Indels 4; Gaps 4;
Db 38 TQNKPLPENVKYGIADAGSSHTNLTYIKWPAEKENDTGVVQQLCECQYKPGISKYAQK 97
0Y 1 TQNKALPENVKYGIADAGSSHTSLTYIKWPAEKENDTGVHVQECCRVKVGPGISKFYQK 60
Db 98 TDEIGAYIAECMESTELIPTSKHHQRPWYIGATAGNRILRMESEQADRYLAAYVTSLEK 157
0Y 61 VNEIGIYLTDMERARREYTPRSQDEPTVYIGATAGNRILRMESEELADRYLDIVERSLS 120

```

```

Db 158 SYPEFQAKITGOERGAYGWITINYLGRPTQEOQSWLSLIS-DSOKOETFGALDIGA 216
QY 121 NTPDFQAGAKITGOERGAYGWITINYLGRFSOKTRWFSIVPEITNNQETFGALDIGA 180
Db 217 STQITFVPQNSTIESPENSIQFRLYGEDYTVYTHSFLCYGKDQALMOKLAKDIQVSSGV 276
QY 181 STQITFVPQNOTIESPDNALQFRLYGKDYNVYTHSFLCYGKDQALMOKLAKDIQVANSNEI 240
Db 277 LKDPCEMPEGYEKVVNVSLEYGTCTKREKKLPDQFRIOGTGDYEOCHOSILEFNNSH 336
QY 241 LKDPCEMPEGYKVVNVSLEYGTCTKREKMTLPQOFIQIGNYOQCHOSILEFNNTSY 300
Db 337 CPYSQCAFNGVFLPPLHSGAFSAFVWDFKVKVAKNSVISOEKMTEITKNPCKSWE 396
QY 301 CPYSQCAFNGVFLPPLHSGAFSAFVWDFKVKVAKNSVISOEKMTEITKNPCKSWE 358
Db 397 EFKTSYPSVKEKYLEXCFSGAYILSL-OGYNFTDSSWEQIHFMGRKIKDSNAGWTLGYM 455
QY 359 EFKTSYAGVKEKYLEXCFSGTYILSLLOGYHETADSWEHIFIGKIQGSDAGWTLGYM 418
Db 456 LNLTMIPAEQPLSPPLPHST 476
QY 419 LNLTMIPAEQPLSPPLSHST 439

```

Search completed: Fri May 5 09:50:43 2000  
 Job time : 835 secs.





(TM)

Release 3.1A John F. Collins, Bioinformatics Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Fri May 5 09:23:53 2000; Maspar time 26.04 Seconds  
Tabular output not generated. 503.481 Million cell updates/sec

Title: >US-09-374-586-2  
Description: (1-439) from US09374586.pep  
Perfect Score: 3275  
Sequence: 1 TQNKALPENVKYGLVDAGS.....NLTNNIPAEQPLSTLSHST 439

Scoring table: PAM 150  
Gap 11

Searched: 82229 seqs, 29864866 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot38  
1:swissprot

Statistics: Mean 50.092; Variance 83.190; scale 0.602

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	3275	100.0	510	1	CD39_HUMAN VASCULAR ATP-DIPHOSPHO	0.00e+00
2	2645	80.8	510	1	CD39_MOUSE VASCULAR ATP-DIPHOSPHO	0.00e+00
3	366	11.2	557	1	YB04_CAEEL HYPOTHETICAL 63.1 KD P	2.58e-54
4	360	11.0	630	1	YEL5_YEAST HYPOTHETICAL 71.9 KD P	5.93e-53
5	323	9.9	485	1	YV4E_CAEEL HYPOTHETICAL 54.3 KD P	1.21e-44
6	316	9.6	454	1	APY_SOLTU APYRASE PRECURSOR (EC	4.34e-43
7	280	8.5	455	1	NTPA_PEA NUCLEOSIDE-TRIPHOSPHAT	3.39e-35
8	265	8.1	518	1	GDAL_YEAST GUANOSINE-DIPHOSPHAT	5.76e-32
9	127	3.9	628	1	NTP2_TOXGO PUTATIVE NUCLEOSIDE-TR	4.46e-05
10	125	3.8	592	1	NTP4_TOXGO PUTATIVE NUCLEOSIDE-TR	9.44e-05
11	125	3.8	628	1	NTP1_TOXGO NUCLEOSIDE-TRIPHOSPHAT	9.44e-05
12	117	3.6	528	1	UDBA_HUMAN UDP-GLUCUCONOSYLTRANSFERASE	1.77e-03
13	102	3.1	450	1	TBA3_MAIZE TUBULIN ALPHA-3 CHAIN	2.99e-01
14	102	3.1	494	1	SYH_ORISA HISTIDIN-TRNA SYNTHETA	2.99e-01
15	101	3.1	530	1	UDBF_HUMAN RNA POLYMERASE SIGMA F	4.13e-01
16	101	3.1	571	1	RESD_CHLTR RNA POLYMERASE SIGMA F	2.99e-01
17	103	3.1	685	1	CRPL_PERRM ALLERGEN CR-P1 PRECURS	2.16e-01
18	103	3.1	787	1	RELA_MYCLE PROBABLE GTP PYROPHOSP	2.16e-01
19	102	3.1	790	1	RELA_MYCTU PROBABLE GTP PYROPHOSP	2.99e-01
20	102	3.1	923	1	PMP2_YEAST PERIODIC TRYPTOPHAN PR	2.99e-01
21	101	3.1	2672	1	GCN1_YEAST TRANSLATIONAL ACTIVATO	4.13e-01
22	97	3.0	410	1	EFTU_CODER ELONGATION FACTOR TU (	1.46e+00
23	97	3.0	841	1	RELA_STRAT GTP PYROPHOSPHOKINASE	1.46e+00

RESULT	ID	CD39_HUMAN	STANDARD	PRT	510 AA.	ALIGNMENTS
AC	P49961					
DT	01-OCT-1996	(Rel. 34, Created)				
DT	01-OCT-1996	(Rel. 34, Last sequence update)				
DT	15-DEC-1998	(Rel. 37, last annotation update)				
DE	VASCULAR ATP-DIPHOSPHOHYDROLASE (EC 3.6.1.5) (ATPase) (LYMPHOID CELL ACTIVATION ANTIGEN) (CD39 ANTIGEN).					
GN	CD39.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;					
OC	Eutheria; Primates; Catarrhini; Homiidae; Homo.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE; 95015846.					
RA	MALISEWSKI C.R., DELSPESE G.J.T., SCHOENBORN M.A., ARMITAGE R.J.,					
RA	FANSLAW W.C., NAKAJIMA T., BAKER E., SUTHERLAND G.R., POINTEUX K.,					
RA	BIRKS C., ALPERT A., FRIEND D., GIMPEL S.D., GAYLE R.B. III;					
RT	"The CD39 lymphoid cell activation antigen. Molecular cloning and structural characterization."					
RT	J. Immunol. 153:3574-3583(1994).					
RT	J. Exp. Med. 185:153-163(1997).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE-UMBILICAL VEIN;					
RX	MEDLINE; 97149443.					
RA	ROBSON S.C., KACZMAREK E., SIEGEL J.B., CANDINAS D., KOZIAK K.,					
RA	MILLAN M., HANCOCK W.W., BACH F.H.;					
RT	"Loss of ATP diphosphohydrolase activity with endothelial cell activation."					
RT	J. Exp. Med. 185:153-163(1997).					
RN	[3]					
RP	FUNCTION.					
RA	KACZMAREK E., KOZIAK K., SEVIGNY J., SIEGEL J.B., ANRATHER J.,					
RA	BRANDOUIN A.R., BACH F.H., ROBSON S.C.;					
RT	"Identification and characterization of CD39/Vascular ATP diphosphohydrolase."					
J. Biol. Chem.	271:33116-33122(1996).					
CC	-1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.					
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).					
CC	-1- TISSUE SPECIFICITY: EXPRESSED PRIMARILY ON ACTIVATED LYMPHOID CELLS. ALSO EXPRESSED IN ENDOTHELIAL TISSUES.					
CC	-1- SIMILARITY: BELONGS TO THE GDAL / CD39 NPase FAMILY.					
CC	-1- DATABASE: NAME=PROV; NOTE=CD guide CD39 entry;					
CC	WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd39.htm"					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; S73813; AAB32152.1; -  
 DR EMBL; U87967; AAB47572.1; -  
 DR MIM; 601752; -  
 DR PROSITE; PS01238; GDAL\_CD39\_NTPASE; 1.

DR PFAM; PF01150; GDAL\_CD39; 1.  
 KW Hydroxylase; Transmembrane; Antigen; Glycoprotein.  
 FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 17 38 POTENTIAL.  
 FT DOMAIN 39 478 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 479 499 POTENTIAL.  
 FT DOMAIN 500 510 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 73 73 POTENTIAL.  
 FT CARBOHYD 227 227 POTENTIAL.  
 FT CARBOHYD 292 292 POTENTIAL.  
 FT CARBOHYD 334 334 POTENTIAL.  
 FT CARBOHYD 371 371 POTENTIAL.  
 FT CARBOHYD 457 457 POTENTIAL.  
 SQ SEQUENCE 510 AA; 57964 MW; E403B5C9 CRC32;

Query Match 100.0%; Score 3275; DB 1; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
 Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 38 TONKALPENVKXGIVLDAGSSHTSLYIKWPAKENDGVVHVEECRVKPGISKEVOK 97  
 QY 1 TONKALPENVKXGIVLDAGSSHTSLYIKWPAKENDGVVHVEECRVKPGISKEVOK 60

Db 98 VNEIGIYITDCMERAREVYPRSOHOETPYLGATAGMRLRMESELDRLVLDVVERSL 157  
 QY 61 VNEIGIYITDCMERAREVYPRSOHOETPYLGATAGMRLRMESELDRLVLDVVERSL 120

Db 158 NYPFPOGARITGOEGAGVMTINYLGRFSOKTRMFSIYPERNNOTFGALDLGA 217  
 QY 121 NYPFPOGARITGOEGAGVMTINYLGRFSOKTRMFSIYPERNNOTFGALDLGA 180

Db 218 STQVTFVPONOTIESPDNALQRLYKGDVNYTHSFLCYGKDQALMOKLADIOVASNEI 277  
 QY 181 STQVTFVPONOTIESPDNALQRLYKGDVNYTHSFLCYGKDQALMOKLADIOVASNEI 240

Db 278 LRDPCHFGYKRVVNVSDLYKTPCKRPFEMTLRPOQFELIOGNGNQCHOSILEFNTSY 337  
 QY 241 LRDPCHFGYKRVVNVSDLYKTPCKRPFEMTLRPOQFELIOGNGNQCHOSILEFNTSY 300

Db 338 CPYSOCAENGFLPLPLGDFGAFSAFYPMKFLNLTSEKVSQEKYTEMMKKFCAQPMEEI 397  
 QY 301 CPYSOCAENGFLPLPLGDFGAFSAFYPMKFLNLTSEKVSQEKYTEMMKKFCAQPMEEI 360

Db 398 KTSYAGVKEKYLEYCFSGTYILSLLOGYHFTADSWEHIFHFGIKQSDAGWTILGYMLN 457  
 QY 361 KTSYAGVKEKYLEYCFSGTYILSLLOGYHFTADSWEHIFHFGIKQSDAGWTILGYMLN 420

Db 458 LTNMIPAEQPLSTPLSHST 476  
 QY 421 LTNMIPAEQPLSTPLSHST 439

RESULT 2 STANDARD: PRT: 510 AA.  
 AC P55772;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE VISCULAR ATP-DIPHOSPHOHYDROLASE (EC 3.6.1.5) (ATPDASE) (LYMPHOID CELL  
 DE ACTIVATION ANTIGEN) (CD39 ANTIGEN).  
 GN CD39.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE; 95015846.  
 RA MALISZEWSKI C.R., DELESPASSE G.J.T., SCHOENBORN M.A., ARMITAGE R.J.,  
 RA FANSLAW W.C., NAKAJIMA T., BAKER E., SUTHERLAND G.R., POINDEXTER R.,  
 RA BIRKS C., ALPERT A., FRIEND D., GIMPEL S.D., GAYLE R.B. III;  
 RT "The CD39 lymphoid cell activation antigen. Molecular cloning and  
 RT structural characterization.";  
 RL J. Immunol. 153:3574-3583(1994).

CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O -> AMP + 2 PHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; AF037366; AAB92259.1; -  
 DR MGD; MGI:102805; CD39.  
 DR PROSITE; PS01238; GDAL\_CD39\_NTPASE; 1.  
 DR PFAM; PF01150; GDAL\_CD39; 1.  
 KW Hydroxylase; Transmembrane; Antigen; Glycoprotein.  
 FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 17 38 POTENTIAL.  
 FT DOMAIN 39 478 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 479 499 POTENTIAL.  
 FT DOMAIN 500 510 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 73 73 POTENTIAL.  
 FT CARBOHYD 226 226 POTENTIAL.  
 FT CARBOHYD 291 291 POTENTIAL.  
 FT CARBOHYD 333 333 POTENTIAL.  
 FT CARBOHYD 428 428 POTENTIAL.  
 FT CARBOHYD 457 457 POTENTIAL.  
 SQ SEQUENCE 510 AA; 57205 MW; 0570BEFE CRC32;

Query Match 80.8%; Score 2645; DB 1; Length 510;  
 Best Local Similarity 76.6%; Pred. No. 0.00e+00;  
 Matches 338; Conservative 57; Mismatches 42; Indels 4; Gaps 4;

Db 38 TONKALPENVKXGIVLDAGSSHTSLYIKWPAKENDGVVHVEECRVKPGISKEVOK 97  
 QY 1 TONKALPENVKXGIVLDAGSSHTSLYIKWPAKENDGVVHVEECRVKPGISKEVOK 60

Db 98 TDEIGAYLAECMELSTELIPTSKHQTPLYLGATAGMRLRMESELDRLVLDVVERSL 157  
 QY 61 VNEIGIYITDCMERAREVYPRSOHOETPYLGATAGMRLRMESELDRLVLDVVERSL 120

Db 158 STPDPFQAKITGOEGAGVMTINYLGRFTQEOSWLSIS-DSQKPTFGALDLGA 216  
 QY 121 NYPFPOGARITGOEGAGVMTINYLGRFSOKTRMFSIYPERNNOTFGALDLGA 180

Db 217 STQVTFVPONOTIESPDNALQRLYKGDVNYTHSFLCYGKDQALMOKLADIOVASGCV 276  
 QY 181 STQVTFVPONOTIESPDNALQRLYKGDVNYTHSFLCYGKDQALMOKLADIOVASNEI 240

Db 277 LRDPCHFGYKRVVNVSDLYKTPCKRPFEMTLRPOQFELIOGNGNQCHOSILEFNTSY 336  
 QY 241 LRDPCHFGYKRVVNVSDLYKTPCKRPFEMTLRPOQFELIOGNGNQCHOSILEFNTSY 300

Db 337 CPYSOCAENGFLPLPLGDFGAFSAFYPMKFLNLTSEKVSQEKYTEMMKKFCAQPMEEI 396  
 QY 301 CPYSOCAENGFLPLPLGDFGAFSAFYPMKFLNLTSEKVSQEKYTEMMKKFCAQPMEEI 358

Db 397 EKTSTYAGVKEKYLEYCFSGTYILSLLOGYHFTADSWEHIFHFGIKQSDAGWTILGYM 455  
 QY 359 EKTSTYAGVKEKYLEYCFSGTYILSLLOGYHFTADSWEHIFHFGIKQSDAGWTILGYM 418

[illegible]

ID	RESULT	4	YEAST	STANDARD	PRT	630 AA
AC	01-FEB-1995	(Rel. 31, Created)				
AD	01-FEB-1995	(Rel. 31, Last sequence update)				
DT	01-NOV-1997	(Rel. 35, Last annotation update)				
DE	HYPOTHEICAL	71.9 KD PROTEIN IN PM140-PC2 INTERGENIC REGION.				
GN	YER005W.					
OS	Saccharomyces cerevisiae	(Baker's yeast).				
OC	Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;					
CC	Saccharomycetaceae; Saccharomyces.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN-S288C / AB972;					
RA	DIETRICH F.S., MULLIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,					
RA	AVILES E., BERNO A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J.M.,					
RA	CHONG E., DUNCAN M., GUZMAN E., HARTFELD G., HUNNICKE-SMITH S.,					
RA	HYMAN R., KAYSER A., KOMP C., LASKARSKI D., LEW H., LIN D.,					
RA	MOSLEDLE D., NAKHAA K., NAMATH A., NORGREN R., OEFNER P., OH C.,					
RA	PEREL F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOGEN T., SMITH V.,					
RA	TAYLOR P., WEI Y., YELTON M., BOYSTEIN D., DAVIS R.W.;					
RL	Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.					
CC	-1- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>					
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).					
CC	-----					
DR	EMBL; U18778; AAB64538.1; .					
DR	PROSITE; PS01238; GDA1_CD39_NTPASE; 1.					
DR	PFAM; PF01150; GDA1_CD39; 1.					
KW	Hypothetical protein; Transmembrane; Hydrolase.					
FT	TRANSMEM 501 517 POTENTIAL.					
SO	SEQUENCE 630 AA; 71851 MW; 68CD0D115 CRC32;					
Query Match	11.0%;	Score 360;	DB 1;	Length 630;		
Best Local Similarity	28.4%;	Pred. No. 5,93e-53;				
Matches 106;	Conservative 93;	Mismatches 137;	Indels 37;	Gaps 28;		
Db	52 IHQKDWTFKLNPGLSSEFEKKPODAYKSHIKPLDPANKIIPESHWSCEPVFIQATAGMR 111					
Qy	41 VHQVECECRK-GPGISKFEQVNEI-GIYLDCMERAREVIRPSOHOEPVYLGATAGMR 98					
Db	112 ILRPDIQSSIIIDLGLOGLKHP-AEFLVEDCSAQIYQISETGLGWMGLNTLYGFEND- 169					
Qy	99 LIRMSSE-LARVLVDVERSLSNPP-DEQQA-RITIQEGEAGWITINILGFFSK 155					
Db	170 --Y-N-P-EVSDHPTFGFMDMGASTQIAFPHDSGEIARRRDIATILFRLSVNDLQK 223					
Qy	156 TRMSIIVPEINNOETFGALDLDGASITQVTFPQNT-IESP-DN-ALQF-R-LYKQ-D- 208					
Db	224 MDVEVSTWLGFGANQARRRYLAQLINTLPENT-ND-YENDDEFSTR-NLMDPCMPRGSSSTD 280					
Qy	209 YVVYTHSEFLCYKQDALQMLAKLDIQVANSNELIRDPCHFPGKRVVNSDLKTP-CTKR 267					
Db	281 PEFKPTI--FHLAGSANGEOCKTSIYPLILKMKPDDDECELCNGVHARIPRIDANKFIQT 338					
Qy	268 FEMTLPEPODFEILQIGNVOOCHOSITILENTSY-CPYSOCAENGILPLPLQ-G-D-FGAF 323					
Db	339 SEYWTANDVFKLGE-YNFDKFSKSLREFCSNNTQIILANSDKGYNSIIPNPLKDACF 397					
Qy	324 SAFIYVVK-FLNLITSEKYSQEKVTEMMKFKCAQPMDEI-----KISYAGVKKKYLSEYCF 377					
Db	398 KGNWVNLILHEGF 410					

RESULT	ID	YV4E_CAEEL	STANDARD:	PRT:	485 AA.
AC	018411				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	01-NOV-1997	(Rel. 35, Last annotation update)			
DE	HYPOTHETICAL 54.3 KD PROTEIN C33H5.14 IN CHROMOSOME IV.				
GN	C33H5.14				
OS	Caenorhabditis elegans.				
OC	Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;				
OC	Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BRISTOL N2;				
RA	BRADSHAW H., STELLYES L.;				
RL	Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.				
CC	-1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see <a href="http://www.isb-sdb.ch/announce/">http://www.isb-sdb.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sdb.ch">license@sdb.ch</a> ).				
CC	-----				
DR	EMBL; U41007; AAA82272.1; -				
DR	WORMPEP; C33H5.14; CE04157.				
DR	PROSITE; PS01238; GDAL_CD39_NTPASE; 1.				
DR	PFAM; PF01150; GDAL_CD39; 1				
KM	Hypothetical protein; Transmembrane; Hydrolase.				
KM	TRANSMEM 439 459 POTENTIAL.				
FT	SEQUENCE 485 AA; 54309 MW; 00659F2B CRC32;				
SO	-----				
Query Match	9.9%; Score 323; DB 1; Length 485;				
Best Local Similarity	25.8%; Pred. No. 1,21e-44;				
Matches	109; Conservative 99; Mismatches 179; Indels 35; Gaps 29;				
Db	21 NNKIVGICDAGSSGTRLEVYTLKPLSGGLNIDPLIHSEFPVKKVYPGLSSFGDKPEQ 80				
QY	8 ENVKGIYLDAGSSHTSLYIKW-P-AEKENDTG-VVHQVECRVK-GPGISKFYQKYN 63				
Db	81 VVEVTLPLRFAEHIPEVQGLGEPOLLIFATGAMLLPEAQDAILIKNLQNLKSGVTALR 140				
QY	64 IGIYITDCMERREVEIPNSQHOETVEYIGATGMLLRNSESSELDRLVDYVERLSNYP 123				
Db	141 VSDSNIRITIDGAMEGIYSMAVNYILGRFDKENDS-KYGMIDMGASVAYAFELINE-KE 198				
QY	124 FDFGQARILITGEBEAYCGMITINILNGKFSQKTRMFSLVPEYETNNQETFGALDLCGASQ 183				
Db	199 -SYNGCN-VYEINLGSIETNEDY-K-YKITYSTELGYCANESL-KKYENSL-VKSGNS-N 251				
QY	184 VTFVPOQNOTIESPDNALQFLR-YGKDYNVYTHSFSLCYGKDAQIMOKLAKDIOVANSNEILR 242				
Db	252 DSCSGRGINRLI--GE-FTVNGTGMDCIA-QVSSL--IGD-KA-QPSC-P--NPT-CF 299				
QY	243 DPCFHPGKTKVANNVSDLLKTPCTKRFEMTLPPQOEIIGYQOCHOSILELFTNTSTCP 302				
Db	300 LRNVAPSVNLSTVOL-YG-FSEWYTTSNFSGGEYHYO-KFTDEVKRYCQKDNNDIOD 356				
QY	303 YSQCAFNGLIPLPGIDGDFGAFSAFFVVKMLNLSEKYSQKEVTEMMKKFCQAPREI-- 360				
Db	357 GFKRNEFPNADIERLTGNCFKAAWYTSVLHGFN-VDKTKHLFGSVLTKIAGEEMQAL 415				
QY	361 --KTS-YAGVEKEYLSEYCFSGTYILSLDQYHFTADSWEHIIHFIKIGIQSGDAGMTIGY 417				
Db	416 ML 417				
QY	418 ML 419				

	RESULT	6	STANDARD:	PRT:	454 AA.
ID	APY_SQUTU				
AC	P80595: Q43164:				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	APYRASE PRECURSOR (EC 3.6.1.5) (ATP-DIPHOSPHATASE) (ADENOSINE DIPHOSPHATASE) (ADPASE) (ATP-DIPHOSPHOHYDROLASE).				
CN	RROPL				
OS	Solanum tuberosum (Potato).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	euphyllophyes; Spermatophyta; Magnoliophyta; eudicotyledons;				
OC	core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;				
CC	Solanum.				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 59-160; 236-253 AND 332-345.				
RC	TISSUE-TUBER:				
KX	MEDLINE; 96158985.				
RA	HANDA M., GUIDOTTI G.;				
RT	"Purification and cloning of a soluble ATP-diphosphohydrolase (apyrase) from potato tubers (Solanum tuberosum).";				
RL	Biochem. Biophys. Res. Commun. 218:916-923(1996).				
RZ	121				
RP	SEQUENCE OF 42-54; 68-95 AND 236-253.				
RC	STRAIN-CV. DESTREE;				
RX	MEDLINE; 96355615.				
RA	KESCONCELOS E.G., FERREIRA S.T., DE CARVALHO T.M.U., DE SOUZA W., VASCONCELOS A.M., MANCILLA M., VALENZUELA M.A., VERJOVSKI-ALMEIDA S.;				
RT	"Partial purification and immunohistochemical localization of ATP diphosphohydrolase from Schistosoma mansoni. Immunological cross-				
RT	reactivities with potato apyrase and Toxoplasma gondii nucleoside triphosphate hydrolase.";				
RL	J. Biol. Chem. 271:22139-22145(1996).				
CC	-1- FUNCTION: CATALYZES THE HYDROLYSIS OF PHOSPHOANHYDRIDE BONDS OF NUCLEOSIDE TRI- AND DI-PHOSPHATES.				
CC	-1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.				
CC	-1- COFACTOR: CALCIUM.				
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED (PROBABLE).				
CC	-1- PTM: THE N-TERMINUS IS BLOCKED.				
CC	-1- SIMILARITY: BELONGS TO THE GAL / CD39 NPYPASE FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on way use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/or send an email to license@isb-sib.ch).				
CC					
DR	EMBL: U58597: AAB02720.1; -				
DR	PROSITE: PS01238; GDALCD39.NPYPASE. 1.				
DR	PFAM: PF01150; GDAL-CD39; 1.				
FT	Hydrolase; Transmembrane; Calcium; Signal.				
FT	FT SIGNAL 1 30 POTENTIAL.				
FT	CHAIN 31 454 APYRASE.				
FT	TRANSMEM 426 446 POTENTIAL.				
FT	CARBOHYD 151 151 POTENTIAL.				
FT	CARBOHYD 262 262 POTENTIAL.				
FT	SEQUENCE 454 AA; 50041 MW; D6FAE89 CRC32;				
DQ	Query Match 9.6%; Score 316; DB 1; Length 454;				
DQ	Best Local Similarity 29.0%; Pred. No. 4,34e-43;				
DQ	Matches 62; Conservative 56; Mismatches 86; Indels 8; Gaps 7;				
DQ	42 ESEHYAVIFDAGSGTSGRVHFRR-DEKLGLLPIGNNIETFMATEPGLSYAEDPKAAANS 100				
OY	8 ENMYKYGIVLDAGSSHSLSLYIKWPAKEKENDGVGHVVEECKRGPGISKFKVQRVNEIGTY 67				
DQ	101 LEPLDGAEBSYVOGDELSETPLELATGTGLMTGDGAEXKLQAVRWLV-KNSTFMSHD 159				
OY	68 LTIDMKAREKVIRDSQHQLTPYLQATGMRLLRHESEE-LADVVLDDVERSLSNTPDF 126				

DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE GANOSINE-DIPHOSPHATASE [EC 3.6.1.42] (GDPase).  
GN GDAL OR YEL042W OR SYGP-ORF16.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
OC Saccharomycetaceae; Saccharomyces.

[1]  
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RP STRAIN-G2-9;  
RX MEDLINE: 93308137.  
RA ABRIOTON C., YANGSISAMA K., MANDON E.C., HARUSHER A., MOREMEN K.,  
RA HIRSCHBERG C.B., ROBBINS P.W.;  
RT "Guanosine diphosphatase is required for protein and sphingolipid  
RL glycosylation in the Golgi lumen of Saccharomyces cerevisiae."  
RM J. Cell Biol. 122:307-323(1993).

[2]  
RN SEQUENCE FROM N.A.  
RP MULIGAN J.T., DIETRICH F.S., HENNESSEY K.M., SEHL P., KOMP C.,  
RA WEI Y., TAYLOR P., NAKAHARA K., ROBERTS D., DAVIS R.W.;  
RL Submitted (FEB-1993) to the EMBL/GeneBank/DDBJ databases.  
RS [3]

RN SEQUENCE FROM N.A.  
RP STRAIN-S28BC / AB972;  
RX DIETRICH F.S., MULIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,  
RA AVYLES E., BERRO A., BENNNAN T., CARPENTER J.J., CHEN E., CHERRY J.M.,  
RA HWANG E., DUCAN M., GUZMAN E., HARTZELL G., HUMICKE-SMITH S.,  
RA HUONG R., KAUSER A., KOMP C., LASHKARI D., LEH H., LIN D.,  
RA MESELDRE D., NAKAHARA K., NAMATH A., NORGEN R., OEPER P., OH C.,  
RA PETREL F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOGREN T., SMITH V.,  
RA TAYLOR P., WEI Y., YELTON M., BORSTEIN D., DAVIS R.W.;  
RL Submitted (DEC-1994) to the EMBL/GeneBank/DDBJ databases.  
RS -1 FUNCTION: AFTER TRANSFER OF SUGARS TO ENDOGENOUS MACROMOLECULAR  
CC ACCEPTORS, THE ENZYME CONVERTS NUCLEOSIDE DIPHSOPHATES TO  
CC NICLOSIDE MONOPHOSPHATES WHICH IN TURN EXIT THE GOLGI LUMEN IN  
CC A COUPLED ANTIPORTER REACTION, ALLOWING ENTRY OF ADDITIONAL  
CC NOCLTOTIDE SUGAR FROM THE CYTOSOL.  
CC -1 CAPATALYTIC ACTIVITY: GDP + H(2)O = GMP + ORTHOPHOSPATE.  
CC CC -1 PATHWAY: GLYCOSTATION.  
CC -1 SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN, GOLGI.  
CC -1 SIMILARITY: BELONGS TO THE GDAL / CD39 NTTPASE FAMILY.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outpostion -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).

-----

DR EMBL: LI9560; AAA34656.1; -;  
DR EMBL: U18779; AAB65000.1; -;  
DR PIR: S30837; S30837.  
DR PIR: AA0732; AA0732.  
DR SGD: L0000695; GDAL.  
DR PROSITE: PS01238; GDAL\_CD39\_NTPASE; 1.  
DR Pfam: PF01150; GDAL\_CD39\_1.  
KW Hydrolase; Golgi stack; Glycoprotein; Transmembrane; Signal-anchor.  
KM DOMAIN 1  
FT DOMAINEM 10 2  
FT DOWNIN 25 518 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN),  
FT DOWNIN 41 41 LUMINAL (POTENTIAL).  
FT CARBOHYD 280 POTENTIAL.  
FT CARBOHYD 335 POTENTIAL.  
SQ SEQUENCE 518 AA; 56821 MW; D139338C CRC32;

Query Match 8.1%; Score 265; DB 1; Length 518;  
Best local similarity 27.8%; Pred. No. 5,76e-32;  
Matches 122; Conservative 92; Mismatches 180; Indels 45; Gaps 37;

0Y PENYKGVIVLDAGSSHTSLYIKKMPAEKENDTVGHQVEQVAKGPISAF-VQKXNIEG 65  
 1 DB 145 S-LDPLLVANMYPIKARCTPVAVKATAGLRLLGPAKSSKILSAVRHLEKDYF-FPY 202  
 2 QY 66 IYLDCCMERAREVYIPRSOHOETPYLGAATGMRL--KMESEELADRVLDYVERSLSNYPF 124  
 3 DB 203 VEGDGVSLMGDEDEGVFMITNTNLLNIGANGP--KL-P--TAA--VF--DLGGSGTP 252  
 4 QY 125 -DFGARILITIOEBGAYMTIITNLLGFSKKTWEISTVPELTNNQETFGALDVGASTQ 183  
 5 DB 253 IVFEETPEINERKMWDEHKRFDLKEDENTLYDPSHLGYGLKES-RKKV-NSVLVE-NA- 308  
 6 QY 184 VTFVPOQNTIHS-PRNALQFRLL-YGKD-YANYTHSFICYGKDQALMQKLAKDIQVANSNEI 240  
 7 DB 309 LKDKIILKGDNTKTHQLSSPCLPPKRVNATNEKVTLESEKTYTIDFGPDESSAGCRFLT 368  
 8 QY 241 LRDP-CFHPGVKRYVNVSDLYKTP-CTKRE-MTLPRQO--FEIIOGNYQOC-HQS-IL- 293  
 9 DB 369 DELUKKDAQOSPSCFENGVQPSLVYTFKESNDIYIEFYDTRTRPLGMLSFTEINELN 428  
 10 QY 294 -ELFMT--SYCYSCQAEFGITLPLQGDV--GA-FSAF-YFVMEFLMT--SEKVSQEKVT 346  
 11 DB 429 DLARIYCKGEETSNVSFSGIAGLDESDSHFCLDLSFQVSLHGTYPIDLP-RE-LRT 486  
 12 QY 347 EMMKFC-AQ-PWEIKRSTYAGVEXKTLSE-YQFSOTYILSLGCGHYHTDSMEHIF 402  
 13 DB 487 GKTIANKKEIGWCLASLPL 505  
 14 QY 403 IGKIOGSDAGWTLTGMLNL 421  
 15 RESULT 9 STANDARD; PRT; 628 AA.  
 16 ID NP2-TOXCO  
 17 IC Q27895; Q2798; Q27801;  
 18 DT 01-NOV-1997 (Rel. 35, Created)  
 19 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 20 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 21 DE NUCLEOSIDE-TRIPHOSPHATASE II PRECURSOR (EC 3.6.1.15) (NTPASE-II)  
 22 DE (NUCLEOSIDE TRIPHOSPHATE HYDROLASE 2).  
 23 GN NP1.  
 24 OS Tokoplasma gondii.  
 25 OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriidae; Sarcocystidae;  
 26 CC Tokoplasma.  
 27 RN [1]  
 28 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.  
 29 RC STRAIN-RH, AND BEVERLEY;  
 30 RX MEDLINE; 95263459.  
 31 RA ASAI T., MUDA S., SIBLEY L.D., OKABAYASHI H., TAKEUCHI T.;  
 32 RT "Biochemical and molecular characterization of nucleoside  
 33 triphosphate hydrolase isozymes from the parasitic protozoan  
 34 Tokoplasma gondii.";  
 35 RT J. Biol. Chem. 270:11391-11397(1995).  
 36 RN [2]  
 37 RP SEQUENCE FROM N.A.  
 38 RC STRAIN-RH;  
 39 RX MEDLINE; 95050750.  
 40 RA BERMEDES D., PECK K.R., AFIFI M.A., BECKERS C.J.M., JOINER K.A.;  
 41 RT "Randomly repeated genes encode nucleoside triphosphate hydrolase  
 42 isoforms secreted into the parasitophorous vacuole of Tokoplasma  
 43 gondii.";  
 44 RT J. Biol. Chem. 269:29252-29260(1994).  
 45 RN [3]  
 46 RP SEQUENCE OF 334-523 FROM N.A.  
 47 RC STRAIN-RH;  
 48 RX MEDLINE; 90152366.  
 49 RA JOHNSON A.M., ILLANA S., MCDONALD P.J., ASAI T.;  
 50 RT "Cloning, expression and nucleotide sequence of the gene fragment  
 51 encoding an antigenic portion of the nucleoside triphosphate  
 52 hydrolase of Tokoplasma gondii.";  
 53 RT Gene 85:215-220(1989).  
 54 CC -I- FUNCTION: MAY PERFORM AN IMPORTANT PROCESSING STEP IN THE  
 55 CONVERSION OF HIGH ENERGY NUCLEOTIDES PRIOR TO UPTAKE BY THE  
 56 PARASITE. NTPASE-II HAS A SPECIFIC ACTIVITY 4.5-FOLD LOWER THAN

```

CC CC NPASE-I IN HYDROLYSES OF ATP. THE PRIMARY DIFFERENCE BETWEEN
CC CC THESE ISOZYMS LIES IN THEIR ABILITY TO HYDROLYSE NUCLEOSIDE
CC CC TRIPHOSPHATE VERSUS DIPHOSPHATE SUBSTRATES. WHILE NPASE-II
CC CC HYDROLYSES ADP TO ADP AND ADP TO AMP AT ALMOST THE SAME RATE,
CC CC NPASE-I HYDROLYSES ADP TO AMP AT A MUCH SLOWER RATE (0.7% OF THE
CC CC RATE FOR ATP).
CC CC
CC CC -I- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + PHOSPHATE.
CC CC -I- SUBUNIT: HOMOTETRAMER.
CC CC -I- SUBCELLULAR LOCATION: SECRETED; FOUND IN HOST CELL PARASITOPHOUS
CC CC VACUOLE.
CC CC
CC CC -I- SIMILARITY: BELONGS TO THE GDA1 / CD39 NPASE FAMILY.
CC CC
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; L39077; AAC41569.1; -.
CC CC DR EMBL; L39079; AAC41570.1; -.
CC CC DR EMBL; U96965; AAC80187.1; -.
CC CC DR EMBL; M33472; AAA30143.1; -.
CC CC DR PROSITE; PS01238; GDA1_CD39_NPASE; 1.
CC CC DR PFAM; PF01150; GDA1_CD39; 1.
CC CC KM Hydrolase; Multigene family; Signal.
CC CC FT SIGNAL 1 25
CC CC FT CHAIN 26 628 NUCLEOSIDE-TRIPHOSPHATASE II.
CC CC FT CARBOHD 432 432 POTENTIAL.
CC CC FT VARIANT 6 6 Y -> H (IN STRAIN BEVERLEY).
CC CC FT VARIANT 91 91 K -> R (IN STRAIN BEVERLEY).
CC CC FT VARIANT 101 101 Q -> R (IN STRAIN BEVERLEY).
CC CC FT CONFLICT 334 334 G -> R (IN REF. 3).
CC CC FT CONFLICT 370 370 V -> L (IN REF. 3).
CC CC FT CONFLICT 372 372 V -> L (IN REF. 3).
CC CC FT CONFLICT 399 399 H -> N (IN REF. 3).
CC CC FT CONFLICT 437 437 E -> K (IN REF. 3).
CC CC FT CONFLICT 488 489 FI -> IV (IN REF. 3).
CC CC FT CONFLICT 492 494 REM -> GGS (IN REF. 3).
CC CC FT CONFLICT 497 497 S -> A (IN REF. 3).
CC CC FT CONFLICT 499 499 D -> N (IN REF. 3).
CC CC FT CONFLICT 523 523 R -> G (IN REF. 3).
CC CC SQ SEQUENCE 628 AA; 69585 MW; DABA65D CRC32;.
CC CC
CC CC Query Match 3.9%; Score 127; DB 1; Length 628;
CC CC Best Local Similarity 32.8%; Pred. No. 4,46e-05;
CC CC Matches 22; Conservative 21; Mismatches 22; Indels 2; Gaps 2;
CC CC
CC CC Db 228 TRPITGAEGLFAETLTNHLNLSRLRGEDPARCMIDYGVKHCNRDIAGVYEWGASAOIVF 287
CC CC :|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
CC CC QY 129 ARLTGOEGAGGAWITIVLLGKFSQKTRMFISIVYETNN-QETP-GALDIGASTQYTF 186
CC CC
CC CC Db 288 PUGECTV 294
CC CC :|:|:|
CC CC QY 187 VPONQTI 193
CC CC
CC CC RESULT 10
CC CC ID NTP4_TOXGO STANDARD: PRT: 592 AA.
CC CC AC P52913.
CC CC DT 01-OCT-1996 (Rel. 34, Created)
CC CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC CC DT 15-DEC-1999 (Rel. 39, Last annotation update)
CC CC DE PURATIVE NUCLEOSIDE-TRIPHOSPHATASE (EC 3.6.1.15) (NUCLEOSIDE
CC CC DE TRIPHOSPHATE PHOSPHOHYDROLASE) (NPASE).
CC CC GN NTP4.
CC CC OS Toxoplasma gondii.
CC CC CC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;
CC CC CC Toxoplasma.
CC CC RN [1]
CC CC SEQUENCE FROM N.A.
CC CC STRAIN-NICOLE;

```

RA COTE S., MORENCY M., LEVESQUE R.C.;  
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + PHOSPHATE.  
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.  
CC -1- CAUTION: THIS COULD BE A PSEUDOGENE. IT DOES NOT SEEM TO START  
WITH AN INITIATOR MET.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; U28353; AAA8036.1; ALT\_INIT.  
CC DR PROSITE; PS01238; GDAL\_CD39\_NTPASE; 1.  
CC DR PFWAM; PF01150; GDAL\_CD39; 1.  
CC KW Hydrolase; Multigene family.  
CC SEQUENCE 592 AA; 65762 MW; DDD65858 CRC32;  
-----  
Query Match 3.8%; Score 125; DB 1; Length 592;  
Best Local Similarity 32.8%; Pred. No. 9,44e-05;  
Matches 22; Conservative 20; Mismatches 23; Indels 2; Gaps 2;  
-----  
DB 192 TRPTGAEGLFAFTTLNHLNRIGEDPARCMIDYGVKQCRNDLAVGVGGASAOIVF 251  
QY 129 ARITIGEGAGVWTTITVLLGKFSQKTRWFSIVYEYETNN-QEYF-GALDVGASTOYTF 186  
DB 252 PLOEGTV 258  
QY 187 VPQNGTI 193  
-----  
RESULT 11  
ID NTP1\_TOXGO STANDARD; PRT; 628 AA.  
AC 027893;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE (NUCLEOSIDE-TRIPHOSPHATASE I PRECURSOR (EC 3.6.1.15) (NTPASE-I)  
DE (NUCLEOSIDE TRIPHOSPHATE HYDROLASE 1).  
GN NTP3.  
OS Toxoplasma gondii.  
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;  
OC Toxoplasma.  
OC [1]  
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.  
RC STRAIN-RH;  
RX MEDLINE; 95263459.  
RA ASAI T., MIDRA S., SIBLEY L.D., OKABAYASHI H., TAKEUCHI T.;  
RT "Biochemical and molecular characterization of nucleoside  
RT triphosphate hydrolase isozymes from the parasitic protozoan  
RT Toxoplasma gondii.";  
RL J. Biol. Chem. 270:11391-11397(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RH;  
RX MEDLINE; 95050750.  
RA BERNHEDS D., PECK K.R., AFFI M.A., BECKERS C.J.M., JOINER K.A.;  
RT "Randomly repeated genes encode nucleoside triphosphate hydrolase  
RT isoforms secreted into the parasitophorous vacuole of Toxoplasma  
RT gondii.";  
RL J. Biol. Chem. 269:29252-29260(1994).  
CC -1- FUNCTION: MAY PERFORM AN IMPORTANT PROCESSING STEP IN THE  
CC CONVERSION OF HIGH ENERGY NUCLEOTIDES PRIOR TO UPTAKE BY THE  
CC PARASITE AND MAY CONTRIBUTE TO INTRACELLULAR SURVIVAL AND  
CC VIRULENCE. NTPASE-I HAS A SPECIFIC ACTIVITY 4.5-FOLD HIGHER THAN  
CC NTPASE-II IN HYDROLYSIS OF ATP. THE PRIMARY DIFFERENCE BETWEEN  
CC THESE ISOZYMES LIES IN THEIR ABILITY TO HYDROLYZE NUCLEOSIDE  
CC TRIPHOSPHATE VERSUS DIPHOSPHATE SUBSTRATES. WHILE NTPASE-II  
CC HYDROLYZES ATP TO ADP AND AMP AT ALMOST THE SAME RATE,  
CC NTPASE-I HYDROLYZES ADP TO AMP AT A MUCH SLOWER RATE (0.7% OF THE

CC RATE FOR ATP).  
CC -1- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + PHOSPHATE.  
CC -1- SUBUNIT: HOMOTETRAMER.  
CC -1- SUBCELLULAR LOCATION: SECRETED; FOUND IN HOST CELL PARASITOPHOUS  
CC VACUOLE.  
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; L39078; AAA89203.1; -;  
CC DR EMBL; U96965; AAC80188.1; -;  
CC DR PROSITE; PS01238; GDAL\_CD39\_NTPASE; 1.  
CC DR PFWAM; PF01150; GDAL\_CD39; 1.  
CC KW Hydrolase; Multigene family; Signal.  
CC FT SIGNAL 1 25  
CC FT CHAIN 26 628 NUCLEOSIDE-TRIPHOSPHATASE I.  
CC FT CARBOHYD 432 432 POTENTIAL.  
CC SEQUENCE 628 AA; 69159 MW; DA2A1577 CRC32;  
-----  
Query Match 3.8%; Score 125; DB 1; Length 628;  
Best Local Similarity 32.8%; Pred. No. 9,44e-05;  
Matches 22; Conservative 20; Mismatches 23; Indels 2; Gaps 2;  
-----  
DB 228 TRPTGAEGLFAFTTLNHLNRIGEDPARCMIDYGVKQCRNDLAVGVGGASAOIVF 287  
QY 129 ARITIGEGAGVWTTITVLLGKFSQKTRWFSIVYEYETNN-QEYF-GALDVGASTOYTF 186  
DB 288 PLOEGTV 294  
QY 187 VPQNGTI 193  
-----  
RESULT 12  
ID UDBA\_HUMAN STANDARD; PRT; 528 AA.  
AC P36537;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE UDP-GLUCURONOSYLTTRANSFERASE 2B10 PRECURSOR, MICROSOAL (EC 2.4.1.17)  
DE (UDPGR).  
GN UGT2B10.  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eukarya; Primates; Catarrhini; Homnidae; Homo.  
OC [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-LIVER;  
RX MEDLINE; 93326164.  
RA JIN C.-J., MINERS J.O., LILLYWHITE K.J., MACKENZIE P.I.;  
RT "cDNA cloning and expression of two new members of the human liver  
RT UDP-glucuronosyltransferase 2B subfamily.";  
RL Biochem. Biophys. Res. Commun. 194:496-503(1993).  
CC -1- FUNCTION: UDPGR IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND  
CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND  
CC ENDOGENOUS COMPOUNDS.  
CC -1- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR  
CC BETA-D-GLUCURONOSIDE.  
CC -1- SUBCELLULAR LOCATION: MICROSOAL.  
CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTTRANSFERASE FAMILY.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



```

CC
DR EMBL: X63359; CAA44961.1; -.
DR PIR: JN0620; JN0620.
DR MIM: 600070; -.
DR PROSITE: PS00375; UDPGT, 1.
DR PFAM: PF00201; UDPGT, 1.
DR Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane; Signal;
KW Multigene family; Microsome.
FT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 1 24 528 UDP-GLUCURONOSYLTRANSFERASE 2B10.
FT TRANSMEM 492 512 POTENTIAL.
FT CARBOHYD 66 66 POTENTIAL.
FT CARBOHYD 314 314 POTENTIAL.
FT CARBOHYD 481 481 POTENTIAL.
SO SEQUENCE 528 AA; 60774 MW; 638280A6 CRC32;

Query Match 3.6%; Score 117; DB 1; Length 528;
Best Local Similarity 24.7%; Pred. No. 1,77e-03;
Matches 21; Conservative 24; Mismatches 36; Indels 4; Gaps 4;

Db 133 KKLKKLOESRPDIYPADAYLPQGEILAEFNIPV-YSH-SFSPGYSFERRSGF-TRP 189
Qy 266 KRFEMTLPEQOEIIGICNGYQOCHOSILELFTMSYCPQSQAFNGIF-LPIQGDGFGAPS 324
Db 190 PSYPPVPMKSLSDQMTFMRVKMNL 214
Qy 325 AFYFVWKFLNLTSEKVSQEKYTEMM 349

RESULT 13
ID TBA3_MAIZE STANDARD: PRT: 450 AA.
AC P22275.
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUBULIN ALPHA-3 CHAIN.
DE TUBA3 OR TUA3.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
OC Poaceae; Zea.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91078640.
RX MONTOLIU L., PUTIGDOMENECH P., RIGAU J.;
RT "The tub alpha 3 gene from Zea mays: structure and expression in
RT dividing plant tissues.";
RL Gene 94:201-207(1990).
RN [2]
RP SEQUENCE FROM N.A.
RP STRAIN-CV. B73; TISSUE-SHOOT;
RX MEDLINE: 92395680.
RX VILLEMER R., JOYE C.M., HAAS N.A., GODDARD R.H., KOPECZAK S.D.,
RA HUSSEY P.J., SNUSTAD D.P., SIEFEL C.D.;
RT "Alpha-tubulin gene family of maize (Zea mays L.). Evidence for two
RT ancient alpha-tubulin genes in plants.";
RL J. Mol. Biol. 227:81-96(1992).
CC -!- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP. ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -!- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -!- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
EMBL: M60171; AAA33518.1; -.
DR EMBL: X63176; CAA44861.1; -.
DR

```

DR	PTR:	JN0105;	JN0105.
DR	PTR:	S28979;	S28979.
DR	MAZEDB:	12141;	-
DR	MENDEL:	1410;	ZEMa;Tuda.3.
DR	PROSITE:	PS00227;	TUBULIN; 1.
DR	PFAM:	PF00091;	tubulin; 1.
KW	Microtubules:	GTP-binding;	Multigene family.
NP	BIND	142	148
FT	SITE	450	450
SO	SEQUENCE	450 AA;	49561 MW; 6E7B522 CRC32; INVOLVED IN POLYMERIZATION.
<hr/>			
Query Match			
Best local Similarity 3.1%; Score 102; DB 1; Length 450;			
Matches 19; Conservative 24; Mismatches 25; Indels 4; Gaps 4;			
Dd	43	GHNDASTFSTFGTGAKHVPRAIVDELPYIDV-RT-GTYQLPHNPQLISGEDA	100
Oy	82	SOHER-PVYLGAAT-GKMRLRSESELADRVLDVERSLSNYPDQGARITLGOEGA	139
<hr/>			
Dd	101	NNFARGHYTKG	112
Oy	140	YGWTITNYLLGK	151
<hr/>			
ID	14	SYH OLYSA	STANDARD; PRT; 494 AA.
AC	P93422;		
Df	15-DEC-1998	(Rel. 37, Created)	
Df	15-DEC-1998	(Rel. 37, Last sequence update)	
Df	15-DEC-1998	(Rel. 37, Last annotation update)	
De	HISTIDYL-TRNA SYNTHETASE (EC 6.1.1.21) (HISTIDINE--TRNA LIGASE)		
De	(HISR).		
Os	Oryza sativa (Rice).		
Oc	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;		
Oc	euphyllophytes: Spermatophyta; Magnoliophyta; Liliopsida; Poales;		
Oc	Poaceae; Oryza.		
Rn	[1]		
Rp	SEQUENCE FROM N.A.		
Rc	STRAIN-CV. NIPPONBARE;		
Ra	AKASHI K., SMALL I.D.;		
Rt	"A cDNA clone encoding rice histidyl-tRNA synthetase."		
Rl	(1) Plant Gene Register PR97-062.		
Cc	-1- CATALYTIC ACTIVITY: ATP + L-HISTIDINE + TRNA(HIS) = AMP +		
Cc	PYROPHOSPHATE + L-HISTIDYL-TRNA(HIS).		
Cc	-1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.		
Cc	-----		
Cc	This SWISS-PROT entry is copyright. It is produced through a collaboration		
Cc	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
Cc	the European Bioinformatics Institute. There are no restrictions on its		
Cc	use by non-profit institutions as long as its content is in no way		
Cc	modified and this statement is not removed. Usage by and for commercial		
Cc	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> )		
Cc	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> .		
Cc	-----		
Dd	EMBL:	Z65984;	CAB0653.1;
Dd	PROSITE:	PS00119;	AA-TRNA LIGASE_II.1; 1.
Dd	PROSITE:	PS00339;	AA-TRNA LIGASE_II.2; 1.
Kw	PFAM:	PF00587;	TRNA-synt.2b.1.
Rv	Aminoacyl-tRNA synthetase:	Protein biosynthesis; Ligase; ATP-binding.	
So	SEQUENCE	494 AA;	55326 MW; 9EB27043 CRC32;
<hr/>			
Query Match			
Best local Similarity 3.1%; Score 102; DB 1; Length 494;			
Matches 23; Conservative 42; Mismatches 40; Indels 11; Gaps 9;			
Dd	128	CSLRDYLVTFAPAYAVMANNISLSKRQYIAKYRRDNPSKGKRYREFYCCDDIDIAGY-ETM	186
Oy	264	CYRKREKTLPEQGF-DIGIGINYQCCHGS-II--E-LFNNSYCPSQCAN--GITPLPL	316
<hr/>			
Dd	187	EPFEVVKVLTLELDQDIGTEIKLNIRKLIDGLMEIGCVPEPKFYVCSSI-DR	241
Oy	317	QGDPGASAFAYFMKFLNLTS-E-KVSDEKYTEMMKKFCQAQPMWEITITSVAGYKEK	370



RESULT 15  
ID UDPF\_HUMAN STANDARD; PRT: 530 AA.  
AC PS4855; P23765;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE UDP-GLUCURONOSYLTRANSFERASE 2B15 PRECURSOR, MICROSOFT (EC 2.4.1.17)  
DE (UNPRT) (UNPRT-3) (HUG44).  
GN UGT2B15 OR UGT2B8.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-LIVER;  
RA MEDLINE; 95136867.  
RX GREEN M.D., OTURU E.M., TEPHLY T.R.;  
RT "Stable expression of a human liver UDP-glucuronosyltransferase  
RT (UGT2B15) with activity toward steroid and xenobiotic substrates.";  
RL Drug Metab. Dispos. 22:799-805(1994).  
RN [2]  
RP SEQUENCE OF 8-530 FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE-LIVER;  
RX MEDLINE; 90343358.  
RA COFFMAN B.L., TEPHLY T.R., IRSHAD Y.M., GREEN M.D., SMITH C.,  
RA JACKSON M.R., WOOSTER R., BURCHELL B.;  
RT "Characterization and primary sequence of a human hepatic microsomal  
RT estriol UDPglucuronosyltransferase.";  
RL Arch. Biochem. Biophys. 281:170-175(1990).  
RN [3]  
RP VARIANT TYR-85.  
RX MEDLINE; 97439504.  
RA LEVESQUE E., BEAULIEU M., GREEN M.D., TEPHLY T.R., BELANGER A.,  
RA HUN D.W.;  
RT "Isolation and characterization of UGT2B15(Y85): a  
RT UDP-glucuronosyltransferase encoded by a polymorphic gene.";  
RL Pharmacogenetics 7:317-325(1997).  
CC -1- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND  
CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND  
CC ENDOGENOUS COMPOUNDS. THIS ISOZYME DISPLAYS ACTIVITY TOWARD  
CC SEVERAL CLASSES OF XENOBIOTIC SUBSTRATES, INCLUDING SIMPLE  
CC PHENOLIC COMPOUNDS, 7-HYDROXYLATED COCAINE, FLAVONOIDS,  
CC ANTHRAQUINONES, AND CERTAIN DRUGS AND THEIR HYDROXYLATED  
CC METABOLITES. IT ALSO CATALYZES THE GLUCURONIDATION OF ENDOGENOUS  
CC ESTROGENS AND ANDROGENS.  
CC -1- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR  
CC BETA-D-GLUCURONOSIDE.  
CC -1- SUBCELLULAR LOCATION: MICROSOFT.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TISSUES.  
CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.  
CC -1- CAUTION: REF.2 THOUGHT THAT THIS WAS A SEPARATE FORM (UGT2B8). THE  
CC NAME UGT2B8 HAS NOW BEEN REUSED FOR A RAT ENZYME.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U08854; AAC50077.1; -;  
DR EMBL; U06641; AAA83406.1; -;  
DR PIR; S11309; S11309.  
DR MIM; 600069; -;  
DR PROSITE; PS00375; UDPGT; 1.  
DR PRAM; PF00201; UDPGT; 1.  
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane; Signal;  
KW Multigene family; Microsome; Polymorphism.  
FT SIGNAL 1 23  
FT CHAIN 24 530 UDP-GLUCURONOSYLTRANSFERASE 2B15.  
FT TRANSMEM 495 515 POTENTIAL.

FT CARBOHYD 65 65 POTENTIAL.  
FT CARBOHYD 316 316 POTENTIAL.  
FT CARBOHYD 483 483 POTENTIAL.  
FT VARIANT 85 85  
FT  
FT CONFLICT 119 119 /FTID=VAR.007713.  
FT CONFLICT 145 145 E -> A (IN REF. 2).  
FT CONFLICT 150 145 K -> R (IN REF. 2).  
FT CONFLICT 152 155 LADALN -> PGDFVF (IN REF. 2).  
FT CONFLICT 162 162 A -> S (IN REF. 2).  
FT CONFLICT 165 165 F -> L (IN REF. 2).  
FT CONFLICT 170 177 LYSLEFSV -> VYRSRSR (IN REF. 2).  
FT CONFLICT 181 181 F -> I (IN REF. 2).  
FT CONFLICT 203 203 S -> I (IN REF. 2).  
FT CONFLICT 293 293 M -> L (IN REF. 2).  
FT CONFLICT 401 401 H -> D (IN REF. 2).  
FT CONFLICT 443 443 V -> A (IN REF. 2).  
FT CONFLICT 501 501 C -> W (IN REF. 2).  
FT CONFLICT 523 523 T -> K (IN REF. 2).  
SQ SEQUENCE 530 AA; 60961 MW; 78E9198E CRC32;  
Query Match 3.1%; Score 102; DB 1; Length 530;  
Best Local Similarity 21.9%; Pred. No. 2.99e-01;  
Matches 21; Conservative 27; Mismatches 44; Indels 4; Gaps 4;

Db 135 KIKMKLOESKFDVTLADALNPGCELLAEFLNIPF-LTSL-RESVGYTFEKNKGGE-LFP 191  
QY 266 KRFEFTLPFQPFETQIGNYQCHOSTLEFNTSTCPYSQCAFNGIF-LPPLQDGFAGFS 324  
Db 192 PSYPPVWSESDQMIEMERIKMIMHLYPFWFQI 227  
QY 325 AFYFMKFLNLTSEKVSQEKTEMMKRCQAPWEI 360

Search completed: Fri May 5 09:26:47 2000  
Job time : 174 secs.



(TM)

```

FT DOMAIN 38 478 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 479 499 POTENTIAL.
FT DOMAIN 500 511 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 73 73 POTENTIAL.
FT CARBOHYD 226 226 POTENTIAL.
FT CARBOHYD 291 291 POTENTIAL.
FT CARBOHYD 333 333 POTENTIAL.
FT CARBOHYD 374 374 POTENTIAL.
FT CARBOHYD 429 429 POTENTIAL.
FT CARBOHYD 458 458 POTENTIAL.
SQ SEQUENCE 511 AA; 57408 MW; E59D8C9 CRC32;

Query Match
Best Local Similarity 80.1%; Score 2624; DB 11; Length 511;
Matches 328; Conservative 63; Mismatches 47; Indels 3; Gaps 3;

Db 98 TDETAAYLAECMKMSTERIPASKOHPYLGATAGMRLRMESKOSADEVLAASRSIK 157
QY 61 VNEIGIYLTDCMERAREVYIPRSOHOETPYLGATAGMRLRMESBELADRYLDVVERSL 120
Db 158 SYPPDFOGAKITGOEGAGYGMTINYLGRFTOBOSWLNFTS-DQKQATEGALDLGS 216
QY 121 NYPDFOGARITITGOEGAGYGMTINYLGRFSQKTRMFSIYPERENNQETFGALDLGA 180
Db 217 STQYTFVPLNLTAPESLQPRLYGTDTYVTHSFLCYGKQALMOKLADIOVSSGI 276
QY 181 STQYTFVPLNLTAPESLQPRLYGTDTYVTHSFLCYGKQALMOKLADIOVSSGI 240
Db 277 LKDCPEYGYKRVVNSVSELYGPTCKREFEKLRFNOFQVQGTGDEQCHSLKFPNNSH 336
QY 241 LRDCPFHGKRYKVVNSDLKTPCKRREMTLPFOQFETIOGIGNQOCHSLKFPNNSH 300
Db 337 CPYSQCAFNGVFLPPLQSGFSAFYVMDFFKKMANSVSSQEKMEITNPFCSKPE 396
QY 301 CPYSQCAFNGVFLPPLQSGFSAFYVMDFFKKMANSVSSQEKMEITNPFCSKPE 358
Db 397 EVKASYPYKREKYLESECFSGTYILSLQGNFTGTSMDQHFHMKIKDSNAGWTLG 456
QY 359 EIKSYAVGKREKYLESECFSGTYILSLQGNFTGTSMDQHFHMKIKDSNAGWTLG 418
Db 457 LNTNMPAEQPLSPPLSHST 477
QY 419 LNTNMPAEQPLSPPLSHST 439

RESULT 2 PRELIMINARY; PRT; 513 AA.
AC 018956;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE VASCULAR ATP-DIPHOSPHOHYDROLASE (EC 3.6.1.5) (LYMPHOID CELL ACTIVATION
DE ANTIGEN) (CD39 ANTIGEN) (ATPASE) (ECTO-ATPASE) (ECTO-
DE ADPASE) (NUCLEOTIDE PHOSPHOHYDROLASE) (ATP PYROPHOSPHOHYDROLASE).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
RN [1]
RP TISSUE-AORTIC ENDOTHELIUM.
RA CHANG A.S., GARCIA R.L., CHANG S.M., SCHILLING W.P.;
RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: HYDROLASES EXTRACELLULAR ATP AND ADP TO AMP. COULD
CC INHIBIT PLATELET AGGREGATION IN RESPONSE TO ADP, COLLAGEN, AND
CC THROMBIN (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.
CC -1- COFACTOR: CALCIUM ION-AND MAGNESIUM ION-DEPENDENT ACTIVITY (BY
CC SIMILARITY).

```

```

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE GDAI / CD39 NPASE FAMILY.
DR EMBL: AF005940; AAB62382.1; -.
DR PROSITE: PS01238; GDAI_CD39_NTPASE; 1.
DR PFAM: PF01150; GDAI_CD39; 1.
KW Hydroxylase; Transmembrane; Antigen; Glycoprotein.
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 37 POTENTIAL.
FT DOMAIN 38 481 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 482 502 POTENTIAL.
FT DOMAIN 503 513 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 73 73 POTENTIAL.
FT CARBOHYD 227 227 POTENTIAL.
FT CARBOHYD 245 245 POTENTIAL.
FT CARBOHYD 307 307 POTENTIAL.
FT CARBOHYD 336 336 POTENTIAL.
FT CARBOHYD 373 373 POTENTIAL.
FT CARBOHYD 460 460 POTENTIAL.
SQ SEQUENCE 513 AA; 58113 MW; 17735933 CRC32;

Query Match
Best Local Similarity 72.9%; Score 2389; DB 6; Length 513;
Matches 309; Conservative 71; Mismatches 58; Indels 5; Gaps 4;

Db 38 TONKALPENVKYGYLDAGSSHTSLYIKWPAKEKENDGVYQTEESNVKPGISGFPAK 97
QY 1 TONKALPENVKYGYLDAGSSHTSLYIKWPAKEKENDGVYQTEESNVKPGISGFPAK 60
Db 98 VNEINYLTLACERAKVYIPSIQHMETPYLGATAGMRLRMENKQADKILAAVASSIS 157
QY 61 VNEIGIYLTDCMERAREVYIPRSOHOETPYLGATAGMRLRMESBELADRYLDVVERSL 120
Db 158 EYPPDFOGARITITGOEGAGYGMTINYLGRFTOBOSWLNFTS-DQKQATEGALDLGA 217
QY 121 NYPDFOGARITITGOEGAGYGMTINYLGRFSQKTRMFSIYPERENNQETFGALDLGA 180
Db 217 STQYTFVPLNLTAPESLQPRLYGTDTYVTHSFLCYGKQALMOKLADIOVSSGI 276
QY 181 STQYTFVPLNLTAPESLQPRLYGTDTYVTHSFLCYGKQALMOKLADIOVSSGI 240
Db 277 LKDCPEYGYKRVVNSVSELYGPTCKREFEKLRFNOFQVQGTGDEQCHSLKFPNNSH 336
QY 241 LRDCPFHGKRYKVVNSDLKTPCKRREMTLPFOQFETIOGIGNQOCHSLKFPNNSH 300
Db 337 CPYSQCAFNGVFLPPLQSGFSAFYVMDFFKKMANSVSSQEKMEITNPFCSKPE 396
QY 301 CPYSQCAFNGVFLPPLQSGFSAFYVMDFFKKMANSVSSQEKMEITNPFCSKPE 358
Db 397 EVKASYPYKREKYLESECFSGTYILSLQGNFTGTSMDQHFHMKIKDSNAGWTLG 456
QY 359 EIKSYAVGKREKYLESECFSGTYILSLQGNFTGTSMDQHFHMKIKDSNAGWTLG 418
Db 457 LNTNMPAEQPLSPPLSHST 477
QY 419 LNTNMPAEQPLSPPLSHST 439

RESULT 3 PRELIMINARY; PRT; 306 AA.
AC 09Y309;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE ECTO-ATP DIPHOSPHOHYDROLASE II (EC 3.6.1.5).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP TISSUE-PLACENTA;
RA MATSUMOTO M., SAKURAI Y., KOKUBO T., YAGI H., MATSUI T., TITANI K.,
RA FUJIMURA Y., NARITA N.;

```

"The cDNA cloning of human placental ecto-ATP diphosphohydrolases I and II.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ133134; CAB41887.1; -  
 DR PROSITE; PS01238; GDAL\_CD39\_NTPASE; 1.

KW Hydrolase.  
 SO SEQUENCE 306 AA; 34175 MM; 2F33C431 CRC32;

Query Match 52.3%; Score 1712; DB 4; Length 306;  
 Best Local Similarity 95.9%; Pred. NO. 0.00e+00;

Matches 236; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

Db 45 TONALPENNYGIVLAGSSHTSLYIKKPAEKENDGVVHOVEECVKGPSKPVOK 104  
 |||||  
 QY 1 TONALPENNYGIVLDGSSHTSLYIKKPAEKENDGVVHOVEECVKGPSKPVOK 60  
 |||||  
 Db 105 VNEIGIYLTDCMERAREVYIPRSOHOETPVYLGATAGMRLMESEBLADRYLDVERSLS 164  
 |||||  
 QY 61 VNEIGIYLTDCMERAREVYIPRSOHOETPVYLGATAGMRLMESEBLADRYLDVERSLS 120  
 |||||  
 Db 165 NYPPDFOGARITITGOEGAGVMTITNLLGFSQKTRMFSTVPTETNNQETFGALDLCGA 224  
 |||||  
 QY 121 NYPPDFOGARITITGOEGAGVMTITNLLGFSQKTRMFSTVPTETNNQETFGALDLCGA 180  
 |||||  
 Db 225 STQVTFPQONOTIESPDNALQFRLYKNDYNYTTHSFILCYGKDALMOKLADIASITQS 284  
 |||||  
 QY 181 STQVTFPQONOTIESPDNALQFRLYKNDYNYTTHSFILCYGKDALMOKLADIASITQS 240  
 |||||  
 Db 285 -RPAPF 289  
 |||||  
 QY 241 LRDPGF 246

RESULT 4  
 ID 093295 PRELIMINARY; PRT; 493 AA.

AC 093295;  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)  
 DE ECTO-ATP-DIPHOSPHOHYDROLASE.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;  
 OC Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=OVIDUCT;  
 RX MEDLINE; 98298108.  
 RA NAGY A.K., KNOWLES A.F., NAGAMI G.T.;  
 RT "Molecular cloning of the chicken oviduct ecto-ATP-  
 diphosphohydrolase.";  
 RL J. Biol. Chem. 273:16043-16049(1998).  
 DR EMBL; AF041355; AAC26491.1; -  
 DR PROSITE; PS01238; GDAL\_CD39\_NTPASE; 1.  
 DR PFAM; PF01150; GDAL\_CD39; 1.  
 KW Hydrolase.  
 SO SEQUENCE 493 AA; 54034 MM; A6E48E68 CRC32;

Query Match 44.4%; Score 1453; DB 13; Length 493;  
 Best Local Similarity 45.5%; Pred. NO. 0.00e+00;

Matches 194; Conservative 88; Mismatches 134; Indels 10; Gaps 10;

Db 35 LPBGTGKGLVDASTHTALYVOMPADKENGITGVSOVESCITYNGSGTSSYADDPAGAG 94  
 |||||  
 QY 6 LPBGTGKGLVDASTHTALYVOMPADKENGITGVSOVESCITYNGSGTSSYADDPAGAG 65  
 |||||  
 Db 95 ASLRKPCDKAAVAVPVOQOMTPYLGATAGMRLMESEBLADRYLDVERSLS 154  
 |||||  
 QY 66 IYLDCHERAREVYIPRSOHOETPVYLGATAGMRLMESEBLADRYLDVERSLS 125  
 |||||  
 Db 155 FRGAQILTGNEEGSGMTVYLLDTLI-KFS-PA-GKWEHPONTVEVYLGALDLCGASTOI 211  
 |||||  
 QY 126 FQGARITITGOEGAGVMTITNLLGFSQKTRMFSTVPTETNNQETFGALDLCGASTOI 184  
 |||||

Db 212 TFPQ-GLTIEDKNMVSLEFLYGTNYSLYTCYGQIOA-SKRLMALHODGSYVONIS 269  
 |||||  
 QY 185 TFPQNOTIESPDNALQFRLYKNDYNYTTHSFILCYGKDALMOKLAKOI-OVASN-ETLR 242  
 |||||  
 Db 270 HPCYPKGYRRIITTAETIDSCVPTPSMISPAQILTYGTGNPAACPAILKFLNLTGCA 329  
 |||||  
 QY 243 DPCFHPGKRVANNVSDLKTPCTKREFTLFPQGEIIGICIGNYQCHOSILEFTSYSCP 302  
 |||||  
 Db 330 NRTGCFDGVYPPVARGOFPFAGFYTFSLNLTGOO-SLSHVNATVDFCNKMSLVE 388  
 |||||  
 QY 303 YSOCAFNGCIFLPLDGDGFAFAFYVAKFLNLSEKVSQEVTEEMKRCFAQPFEEIKT 362  
 |||||  
 Db 389 TFPQNK-GLTHTYCVGGLYITLLVDYKFKDEHMSNTHFQKAGNADIGTGLNLT 447  
 |||||  
 QY 363 SYAGVKEKYLSECHPSGYTLLSLDQGHYTFADSWERHIFGKIQGSDAGWTLGYMLNT 422  
 |||||

Db 448 NMIPTE 453  
 |||||  
 QY 423 NMIPAE 428

RESULT 5  
 ID 055026 PRELIMINARY; PRT; 495 AA.  
 AC 055026; 035928;  
 DT 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)  
 DE ECTO-ATPASE (EC 3.6.1.5).  
 GN CD39L1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RX MEDLINE; 98288263.  
 RA GAO L., DONG L., WHITLOCK J.P. JR.;  
 RT "A novel response to dioxin. Induction of ecto-ATPase gene  
 expression.";  
 RL J. Biol. Chem. 273:15358-15365(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A. (SHORT FORM).  
 RC TISSUE=EMBRYO;  
 RX MEDLINE; 97419269.  
 RA CHADWICK B.P., FRISCHAUF A.M.;  
 RT "Cloning and mapping of a human and mouse gene with homology to ecto-  
 ATPase genes.";  
 RT Mamm. Genome 8:668-672(1997).

CC -1- FUNCTION: HYDROLYSES EXTRACELLULAR ATP TO AMP.  
 CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- INDUCTION: BY DIOXIN.

CC -1- PTM: HAS PROBABLY VARIOUS DISULFIDE-BONDS (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.

DR EMBL; AF042811; AAC24347.1; -  
 DR EMBL; U91511; AAB81014.1; -  
 DR MGD; MGI:1096863; C63911.  
 DR PROSITE; PS01238; GDAL\_CD39\_NTPASE; 1.  
 DR PFAM; PF01150; GDAL\_CD39; 1.  
 KW Hydrolase; Transmembrane; Antigen; Glycoprotein; Alternative splicing.  
 SO CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 1 4  
 FT TRANSMEM 25  
 FT DOMAIN 22 462  
 FT TRANSMEM 463 483  
 FT DOMAIN 467 470  
 FT DOMAIN 484 495  
 FT VARSPPLIC 130 132  
 FT VARSPPLIC 133 495  
 FT CARBOHYD 64 64  
 FT CARBOHYD 129 129

FT POLY-LBU. EXTRACELLULAR (POTENTIAL).  
 FT POLY-LBU. POTENTIAL.  
 FT CYTOPLASMIC (POTENTIAL).  
 FT LYS -> MAG (IN SHORT FORM).  
 FT MISSING (IN SHORT FORM).  
 FT POTENTIAL.  
 FT POTENTIAL.

Query Match	Best Local Similarity	Score	DB	Length
Matches 199; Conservative	45.5%;	1361;	DB 11;	495;
	83;	Pred. No. 4,398-284;		
		Mismatches 139;		Indels 15; Gaps
DOMAIN	1			
TRANSMEM	5			
FT	25			
DOMAIN	22			
FT	25			
DOMAIN	26			
FT	462			
TRANSMEM	463			
FT	483			
DOMAIN	467			
FT	470			
DOMAIN	484			
FT	495			
CARBOHYD	64			
FT	64			
CARBOHYD	129			
FT	129			
CARBOHYD	294			
FT	294			
CARBOHYD	306			
FT	306			
CARBOHYD	319			
FT	319			
CARBOHYD	378			
FT	378			
CARBOHYD	443			
FT	443			
SEQUENCE	495 AA;	54389 MW;	24f8E16C	CRC32;
Query Match	41.6%;	Score 1361;	DB 11;	Length 495;
Best Local Similarity	45.5%;	Pred. No. 4,398-284;		
Matches 199; Conservative	83;	Mismatches 139;		Indels 15; Gaps
DOMAIN	1			
TRANSMEM	5			
FT	25			
DOMAIN	22			
FT	25			
DOMAIN	26			
FT	462			
TRANSMEM	463			
FT	483			
DOMAIN	467			
FT	470			
DOMAIN	484			
FT	495			
CARBOHYD	64			
FT	64			
CARBOHYD	129			
FT	129			
CARBOHYD	294			
FT	294			
CARBOHYD	306			
FT	306			
CARBOHYD	319			
FT	319			
CARBOHYD	378			
FT	378			
CARBOHYD	443			
FT	443			
SEQUENCE	495 AA;	54389 MW;	24f8E16C	CRC32;
Query Match	41.6%;	Score 1361;	DB 11;	Length 495;
Best Local Similarity	45.5%;	Pred. No. 4,398-284;		
Matches 199; Conservative	83;	Mismatches 139;		Indels 15; Gaps
DOMAIN	1			
TRANSMEM	5			
FT	25			
DOMAIN	22			
FT	25			
DOMAIN	26			
FT	462			
TRANSMEM	463			
FT	483			
DOMAIN	467			
FT	470			
DOMAIN	484			
FT	495			
CARBOHYD	64			
FT	64			
CARBOHYD	129			
FT	129			
CARBOHYD	294			
FT	294			
CARBOHYD	306			
FT	306			
CARBOHYD	319			
FT	319			
CARBOHYD	378			
FT	378			
CARBOHYD	443			
FT	443			
SEQUENCE	495 AA;	54389 MW;	24f8E16C	CRC32;







SEQUENCE	472 AA;	51161 MW;	56924C38 CRC32;
Query Match	33.0%;	Score 1080;	DB 4; Length 472;
Best Local Similarity	43.4%;	Pred. No. 6,40e-217;	
Matches	155;	Conservative	80; Mismatches 110; Indels 12; Gaps 11.
Db	35	PALATYGIIVLADGSSHTSMETFKKPAKDEKNTGYIGQHSQSDVYGGGSSSTADNPPSQAQ	94
Oy	7	PENAYKGVILADGSSHTSLYIKWPAKDEKNTGVHGVHVEEVRVGPISKRVQKVNIGI	66
Db	95	SLVGLDEALADVPERHAGTPLYIGATAGKRLNLTNPEASTVYLAIVHTTLQYEPDF	154
Oy	67	YLTDMEARREYIPRSQOETFPVILGATAGKRLRMESSEELADRYLDVVERSLSNYPDF	126
Db	155	RCARILSGQEBGVETWYANYLLENFI-KYGVWR-WFRPRK-CTLGAMDIGASTOITF	211
Oy	127	OGARITGOEGSAGWITINYLTKGFSQKTFWFSIVPETNNOETFGALDIGASTOITF	186
Db	212	ETTSTA-EDRASVQALHGYOHRYTHSLCTYGRDYL-ORLLASA-LOTGCF-H-PCW	266
Oy	187	VPQNTIESPNNALOFRLYGDYNYTHSFICYKDDALMOKLAKDIQVANSNELLRPOCF	246
Db	267	PRGFTGYLLDDVYQSPPTMAQNPONFNSSARVSLSSGSDPHLCRDYSGIFSSPSQFS	326
Oy	247	HPGKAYVNVSDLLKTPCTK--KREMTLPFOQFEIIGGNTQOCHQSLTEFNTSYCPYS	304
Db	327	RCSENGVQPEVPAGNFVAFSAFETVDFLTRSMGLPYATLQQLAAVAVNCNQTAQ	383
Oy	305	QCAFNGIFLPLQDGFAGFSAFYVMKRLNTLSE-KVSO-EKVTENMKKRCQAQWEE	359
RESULT	12	PRELIMINARY;	PRT; 483 AA.
ID	OQX162	AC	OQX162;
DT	01-NOV-1999	(TREMBLrel, 12, Created)	
DT	01-NOV-1999	(TREMBLrel, 12, Last sequence update)	
DT	01-NOV-1999	(TREMBLrel, 12, Last annotation update)	
DE	F7A19.34	PROTEIN.	
GN	F7A19.34		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;		
OC	core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;		
OC	Arabidopsis.		
RN	[1]		
RP	SEQUENCE FROM N. A.		
RC	STRAIN-CV, COLUMBIA.		
RA	FRKESTEL N. A., PALM C. J., CONWAY A. B., CONN L., HANSEN N. F.,		
RA	ALFAFI H., ABAUO R., HUIZAR L., ROWLEY D., BUEHLER E., DUNN P.,		
RA	GOLOS A. A., KREMETSCHTA I. I., KIM C., LENZ C., LI J., LIU S.,		
RA	LUNOS S., SCHWARTZ J., SHINN P., TORIOWI M., VYSTOSKAIA V. S.,		
RA	WALKER M., YU G., ECKER J., THEOLOGIS A., DAVIS R. W.,		
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AC007576; AAD39311.1;		
SO	SEQUENCE 483 AA; 53425 MW; B134313C CRC32;		
Query Match	18.0%;	Score 590;	DB 10; Length 483;
Best Local Similarity	30.9%;	Pred. No. 4.01e-102;	
Matches	131;	Conservative	104; Mismatches 154; Indels 35; Gaps 25.
Db	67	KLRYSVLIDAGSSGTRVHFGVWFESGKPVDFEGEKHYANL-KLT-PGLSYADNPEGAS	124
Oy	9	NVKKGIVLADGSSHTSLYIK-W-PAEKEN-DTGVHGVHVEEVRVGPISKRVQKVNIGI	65
Db	125	VSVYKLVFAQRIKPKRFRSDIRLMATAGMLL--EVP-VQDQILTEVTRVLRSSGEM	181
Oy	66	YILDQCMERAREVYIPRSQOETFPVILGATAGKRLRMESSEELADRYLDVVERSLSNYPDF	125
Db	182	FRDEMANVSSDGIYSMTANALGSLG--T--D--PLET-----T-GIVELGSGSAQ	229
Oy	126	FOG--ARITITGOEGSAGWITINYLTKGFSQKTFWFSIVPETNNOETFGALDIGASTQ	183
Db	230	VTFVSSSEH-V-PEVYSRTIA-VGNISYTIYSHSFLDYGKDAAL-KLLEKLONSANSTV-	284

QY	14	YTFVFNQNTIESPDNALQFLRGK-DYVVNYTHSEFLCYCKDDALWOKLAKKDIOVASNETLR	242
Db	285	DGVNDDPCTPKGYIYDNTSKNYSNGSGLFADNDESKLKSIOAANFSKRSRSPALKEGEN	344
OY	243	DCGFHDPGRKKVYVNSDYLLKTPCTKRFEMTLPRQOEIOGICIGNYQOCHSIIELFNITS--Y	300
Db	345	CLYEHCSIGSFPTPDLOGSFATASFYTTAFPEEL-EEKGMLSLLIPAGKRYCEEMSKL	403
OY	301	CPYSCAANGFJLPLRQDDEGAFSAFYFVMEKFLNLTSEKVSQEKVETEMKKRFAQMPWEI	360
Db	404	ILEYFTDEBYLRGCFESAATITSLMLHDSLGIALDD-ESTIYASKAGGKHLPLDMALGAF	462
OY	361	KTSYAGVREKYLEXCFSGFYTLISLLLOGYHFTADSWEHIFGRKIOGSD-A-QWTLG-Y	417
Db	463	ILDV 466	
OY	418	MLNL 421	
RESULT	13	PRELIMINARY; PRT; 609 AA.	
ID	015092		
AC	015092		
DT	01-JAN-1998	(TREMBLrel. 05, Created)	
DT	01-NOV-1998	(TREMBLrel. 08, Last sequence update)	
DT	01-NOV-1998	(TREMBLrel. 08, Last annotation update)	
DE	GUANOSINE-DIPHOSPHATASE LIKE PROTEIN (XIA0392).		
GN	XIA0392.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
OC	Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
NC	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-BRAIN:		
RX	MEDLINE; 98225229.		
RA	WANG T.F., GUIDOTTI G.;		
RT	"Golgi localization and functional expression of human uridine		
RT	diphosphatase.";		
RL	J. Biol. Chem. 273:11392-11399(1998).		
RM	[2]		
RP	SEQUENCE OF 60-609 FROM N.A.		
RC	TISSUE-BRAIN:		
RX	MEDLINE; 97349984.		
RA	NAKASE T., ISHIKAWA K., NAKAJIMA D., OHIRA M., SEKI N., MIYAJIMA N.,		
RA	TAUKA A., KOTANI H., NOMURA N., OHARA O.;		
RT	"Prediction of the coding sequences of unidentified human genes. VII.		
RT	The complete sequences of 100 new cDNA clones from brain which can		
RT	code for large proteins in vitro.";		
RL	DNA Res. 4:141-150(1997).		
DR	EMBL; AF016032; AAC17217.1; -		
DR	EMBL; AB002390; BAC21575.1; -		
DR	PFAM; PF01150; GDA1CD39; 1.		
DR	SEQUENCE 609 AA; 69491 MW; 6A366FC5 CRC32;		
QO			
Query Match	12.1%;	Score 397; DB 4; Length 609;	
Best Local Similarity	12.0%;	Pred. No. 1,00e-58;	
Matches	129; Conservative	120; Mismatches 155; Indels 56; Gaps 43	
Db	87	NVNTGIYVDDCGSSGRVETVCPMRHNGPNDLLDIRQNRDKNRKRVYMKIKGISERATS	146
OY	9	NVKYGIYVDDCGSSGRVETVCPMRHNGPNDLLDIRQNRDKNRKRVYMKIKGISERATS	146
Db	147	PEKYSVDSYPLINFAEYVPRAKHKEPPLYLTCRAGMRILP-ESQKA--TLEDLTDIP	203
OY	61	VNEIGITVITDMEKRRREYIPRSQOEFTYVYGAATGAKMLMESEELADRYLDVVERSL	120
Db	204	VHFDLFESDHAERYISGKOEQVYMGINFEVYGRFEHIEDDEAVVEVNIIGSESSEAIY	263
OY	121	NYPDF-QG-ARIITGEBGAYGMITVNLGKFSQ---KTRAF-SI-VP-YETNN---	168
Db	264	RKFRAGIIDMGGVSTQIAYEVKTEEVAKNLLAEFNLCSDVHOTEHRYRVYVATFLGEGG	323
OY	169	OE-FEGADLDIGASTOVWF-VPOQNTIE-S--PDNALQFLRYGKD--YNYTTHSEFLCYGK	221

